VOLUME VI MARCH, 1908 No. 1

# BIOMETRIKA.

## THE PROBABLE ERROR OF A MEAN.

BY STUDENT.

### *Introduction.*

ANY experiment may be regarded as forming an individual of a "population" of experiments which might be performed under the same conditions. A series of experiments is a sample drawn from this population.

Now any series of experiments is only of value in so far as it enables us to form a judgment as to the statistical constants of the population to which the experiments belong. In a great number of cases the question finally turns on the value of a mean, either directly, or as the mean difference between the two quantities.

If the number of experiments be very large, we may have precise information as to the value of the mean, but if our sample be small, we have two sources of uncertainty:—(1) owing to the "error of random sampling" the mean of our series of experiments deviates more or less widely from the mean of the population, and (2) the sample is not sufficiently large to determine what is the law of distribution of individuals. It is usual, however, to assume a normal distribution, because, in a very large number of cases, this gives an approximation so close that a small sample will give no real information as to the manner in which the population deviates from normality: since some law of distribution must be assumed it is better to work with a curve whose area and ordinates are tabled, and whose properties are well known. This assumption is accordingly made in the present paper, so that its conclusions are not strictly applicable to populations known not to be normally distributed; yet it appears probable that the deviation from normality must be very extreme to lead to serious error. We are concerned here solely with the first of these two sources of uncertainty.

The usual method of determining the probability that the mean of the population lies within a given distance of the mean of the sample, is to assume a normal distribution about the mean of the sample with a standard deviation equal to $s/\sqrt{n}$, where $s$ is the standard deviation of the sample, and to use the tables of the probability integral.

But, as we decrease the number of experiments, the value of the standard deviation found from the sample of experiments becomes itself subject to an increasing error, until judgments reached in this way may become altogether misleading.

In routine work there are two ways of dealing with this difficulty: (1) an experiment may be repeated many times, until such a long series is obtained that the standard deviation is determined once and for all with sufficient accuracy. This value can then be used for subsequent shorter series of similar experiments. (2) Where experiments are done in duplicate in the natural course of the work, the mean square of the difference between corresponding pairs is equal to the standard deviation of the population multiplied by $\sqrt{2}$. We can thus combine together several series of experiments for the purpose of determining the standard deviation. Owing however to secular change, the value obtained is nearly always too low, successive experiments being positively correlated.

There are other experiments, however, which cannot easily be repeated very often; in such cases it is sometimes necessary to judge of the certainty of the results from a very small sample, which itself affords the only indication of the variability. Some chemical, many biological, and most agricultural and large scale experiments belong to this class, which has hitherto been almost outside the range of statistical enquiry.

Again, although it is well known that the method of using the normal curve is only trustworthy when the sample is "large," no one has yet told us very clearly where the limit between "large" and "small" samples is to be drawn.

The aim of the present paper is to determine the point at which we may use the tables of the probability integral in judging of the significance of the mean of a series of experiments, and to furnish alternative tables for use when the number of experiments is too few.

The paper is divided into the following nine sections:

I. The equation is determined of the curve which represents the frequency distribution of standard deviations of samples drawn from a normal population.

II. There is shown to be no kind of correlation between the mean and the standard deviation of such a sample.

III. The equation is determined of the curve representing the frequency distribution of a quantity $z$, which is obtained by dividing the distance between the mean of a sample and the mean of the population by the standard deviation of the sample.

IV. The curve found in I. is discussed.

V. The curve found in III. is discussed.

VI. The two curves are compared with some actual distributions.

VII. Tables of the curves found in III. are given for samples of different size.

VIII and IX. The tables are explained and some instances are given of their use.

X. Conclusions.

## Section I.

Samples of $n$ individuals are drawn out of a population distributed normally, to find an equation which shall represent the frequency of the standard deviations of these samples.

If $s$ be the standard deviation found from a sample $x_1 x_2 \ldots x_n$ (all these being measured from the mean of the population), then

$$s^2 = \frac{S(x_1^2)}{n} - \left(\frac{S(x_1)}{n}\right)^2 = \frac{S(x_1^2)}{n} - \frac{S(x_1^2)}{n^2} - \frac{2S(x_1x_2)}{n^2}.$$

Summing for all samples and dividing by the number of samples we get the mean value of $s^2$ which we will write $\bar{s}^2$,

$$\bar{s}^2 = \frac{n\mu_2}{n} - \frac{n\mu_2}{n^2} = \frac{\mu_2(n-1)}{n},$$

where $\mu_2$ is the second moment coefficient in the original normal distribution of $x$: since $x_1$, $x_2$, etc., are not correlated and the distribution is normal, products involving odd powers of $x_1$ vanish on summing, so that $\frac{2S(x_1x_2)}{n^2}$ is equal to 0.

If $M_R'$ represent the $R^{\text{th}}$ moment coefficient of the distribution of $s^2$ about the end of the range where $s^2 = 0$,

$$M_1' = \mu_2 \frac{(n-1)}{n}.$$

Again
$$s^4 = \left\{\frac{S(x_1^2)}{n} - \left(\frac{S(x_1)}{n}\right)^2\right\}^2$$

$$= \left(\frac{S(x_1^2)}{n}\right)^2 - \frac{2S(x_1^2)}{n}\left(\frac{S(x_1)}{n}\right)^2 + \left(\frac{S(x_1)}{n}\right)^4$$

$$= \frac{S(x_1^4)}{n^2} + \frac{2S(x_1^2x_2^2)}{n^2} - \frac{2S(x_1^4)}{n^3} - \frac{4S(x_1^2x_2^2)}{n^3} + \frac{S(x_1^4)}{n^4}$$

$$+ \frac{6S(x_1^2x_2^2)}{n^4} + \text{other terms involving odd powers of } x_1 \text{ etc., which will vanish on summation.}$$

Now $S(x_1^4)$ has $n$ terms but $S(x_1^2x_2^2)$ has $\frac{1}{2}n(n-1)$, hence summing for all samples and dividing by the number of samples we get

$$M_2' = \frac{\mu_4}{n} + \mu_2^2\frac{(n-1)}{n} - \frac{2\mu_4}{n^2} - 2\mu_2^2\frac{(n-1)}{n^2} + \frac{\mu_4}{n^3} + 3\mu_2^2\frac{(n-1)}{n^3}$$

$$= \frac{\mu_4}{n}\{n^2 - 2n - 1\} + \frac{\mu_2^2}{n^3}(n-1)\{n^2 - 2n + 3\}.$$

Now since the distribution of $x$ is normal, $\mu_4 = 3\mu_2^2$, hence

$$M_2' = \mu_2^2\frac{(n-1)}{n^3}\{3n - 3 + n^2 - 2n + 3\} = \mu_2^2\frac{(n-1)(n+1)}{n^2}.$$

In a similar tedious way I find:

$$M_3' = \mu_2^3 \frac{(n-1)(n+1)(n+3)}{n^3},$$

and

$$M_4' = \mu_2^4 \frac{(n-1)(n+1)(n+3)(n+5)}{n^4}.$$

The law of formation of these moment coefficients appears to be a simple one, but I have not seen my way to a general proof.

If now $M_R$ be the $R^{\text{th}}$ moment coefficient of $s^2$ about its mean, we have

$$M_2 = \mu_2^2 \frac{(n-1)}{n^2} \{(n+1)-(n-1)\} = 2\mu_2^2 \frac{(n-1)}{n^2},$$

$$M_3 = \mu_2^3 \left\{ \frac{(n-1)(n+1)(n+3)}{n^3} - 3\frac{(n-1)}{n} \cdot \frac{2(n-1)}{n^2} - \frac{(n-1)^3}{n^3} \right\}$$

$$= \mu_2^3 \frac{(n-1)}{n^3} \{n^2+4n+3-6n+6-n^2+2n-1\} = 8\mu_2^3 \frac{(n-1)}{n^3},$$

$$M_4 = \frac{\mu_2^4}{n^4} \{(n-1)(n+1)(n+3)(n+5) - 32(n-1)^2 - 12(n-1)^3 - (n-1)^4\}$$

$$= \frac{\mu_2^4 (n-1)}{n^4} \{n^3+9n^2+23n+15-32n+32-12n^2+24n-12-n^3+3n^2-3n+1\}$$

$$= \frac{12\mu_2^4 (n-1)(n+3)}{n^4}.$$

Hence

$$\beta_1 = \frac{M_3^2}{M_2^3} = \frac{8}{n-1}, \quad \beta_2 = \frac{M_4}{M_2^2} = \frac{3(n+3)}{n-1},$$

$$\therefore\ 2\beta_2 - 3\beta_1 - 6 = \frac{1}{n-1} \{6(n+3) - 24 - 6(n-1)\} = 0.$$

Consequently a curve of Professor Pearson's type III. may be expected to fit the distribution of $s^2$.

The equation referred to an origin at the zero end of the curve will be

$$y = Cx^p e^{-\gamma x},$$

where

$$\gamma = 2\frac{M_2}{M_3} = \frac{4\mu_2^2 (n-1) n^3}{8n^2 \mu_2^3 (n-1)} = \frac{n}{2\mu_2},$$

and

$$p = \frac{4}{\beta_1} - 1 = \frac{n-1}{2} - 1 = \frac{n-3}{2}.$$

Consequently the equation becomes

$$y = Cx^{\frac{n-3}{2}} e^{-\frac{nx}{2\mu_2}},$$

which will give the distribution of $s^2$.

The area of this curve is $C\int_0^\infty x^{\frac{n-3}{2}} e^{-\frac{nx}{2\mu_2}} dx = I$ (say).

The first moment coefficient about the end of the range will therefore be

$$\frac{C\int_0^\infty x^{\frac{n-1}{2}} e^{-\frac{nx}{2\mu_2}} dx}{I} = \frac{C\left[\frac{-2\mu_2}{n} x^{\frac{n-1}{2}} e^{-\frac{nx}{2\mu_2}}\right]_{x=0}^{x=\infty}}{I} + \frac{C\int_0^\infty \frac{n-1}{n}\mu_2 x^{\frac{n-3}{2}} e^{-\frac{nx}{2\mu_2}} dx}{I}.$$

The first part vanishes at each limit and the second is equal to

$$\frac{\frac{n-1}{n}\mu_2 I}{I} = \frac{n-1}{n}\mu_2,$$

and we see that the higher moment coefficients will be formed by multiplying successively by $\frac{n+1}{n}\mu_2$, $\frac{n+3}{n}\mu_2$, etc., just as appeared to be the law of formation of $M_2'$, $M_3'$, $M_4'$, etc.

Hence it is probable that the curve found represents the theoretical distribution of $s^2$; so that although we have no actual proof we shall assume it to do so in what follows.

The distribution of $s$ may be found from this, since the frequency of $s$ is equal to that of $s^2$ and all that we must do is to compress the base line suitably.

Now if $y_1 = \phi(s^2)$ be the frequency curve of $s^2$

and $y_2 = \psi(s)$ „ „ „ „ „ $s$,

then $y_1 d(s^2) = y_2 ds$,

or $y_2 ds = 2y_1 s ds$,

$\therefore\ y_2 = 2s y_1$.

Hence $y_2 = 2Cs(s^2)^{\frac{n-3}{2}} e^{-\frac{ns^2}{2\mu_2}}$

is the distribution of $s$.

This reduces to $y_2 = 2Cs^{n-2} e^{-\frac{ns^2}{2\mu_2}}$.

Hence $y = Ax^{n-2} e^{-\frac{nx^2}{2\sigma^2}}$ will give the frequency distribution of standard deviations of samples of $n$, taken out of a population distributed normally with standard deviation $\sigma$. The constant $A$ may be found by equating the area of the curve as follows:—

$$\text{Area} = A\int_0^\infty x^{n-2} e^{-\frac{nx^2}{2\sigma^2}} dx. \quad \left(\text{Let } I_p \text{ represent } \int_0^\infty x^p e^{-\frac{nx^2}{2\sigma^2}} dx.\right)$$

Then
$$I_p = \frac{\sigma^2}{n}\int_0^\infty x^{p-1}\frac{d}{dx}\left(-e^{-\frac{nx^2}{2\sigma^2}}\right) dx$$
$$= \frac{\sigma^2}{n}\left[-x^{p-1} e^{-\frac{nx^2}{2\sigma^2}}\right]_{x=0}^{x=\infty} + \frac{\sigma^2}{n}(p-1)\int_0^\infty x^{p-2} e^{-\frac{nx^2}{2\sigma^2}} dx$$
$$= \frac{\sigma^2}{n}(p-1) I_{p-2},$$

since the first part vanishes at both limits.

By continuing this process we find

$$I_{n-2} = \left(\frac{\sigma^2}{n}\right)^{\frac{n-2}{2}} (n-3)(n-5)\dots 3.1\, I_0$$

or
$$= \left(\frac{\sigma^2}{n}\right)^{\frac{n-3}{2}} (n-3)(n-5)\dots 4.2\, I_1$$

according as $n$ is even or odd.

But $I_0$ is
$$\int_0^\infty e^{-\frac{nx^2}{2\sigma^2}} dx = \sqrt{\frac{\pi}{2n}}\,\sigma,$$

and $I_1$ is
$$\int_0^\infty x e^{-\frac{nx^2}{2\sigma^2}} dx = \left[-\frac{\sigma^2}{n} e^{-\frac{nx^2}{2\sigma^2}}\right]_{x=0}^{x=\infty} = \frac{\sigma^2}{n}.$$

Hence if $n$ be even,

$$A = \frac{\text{Area}}{(n-3)(n-5)\dots 3.1 \sqrt{\frac{\pi}{2}} \left(\frac{\sigma^2}{n}\right)^{\frac{n-1}{2}}},$$

and if $n$ be odd

$$A = \frac{\text{Area}}{(n-3)(n-5)\dots 4.2 \left(\frac{\sigma^2}{n}\right)^{\frac{n-1}{2}}}.$$

Hence the equation may be written

$$y = \frac{N}{(n-3)(n-5)\dots 3.1} \sqrt{\frac{2}{\pi}} \left(\frac{n}{\sigma^2}\right)^{\frac{n-1}{2}} x^{n-2} e^{-\frac{nx^2}{2\sigma^2}} \quad (n \text{ even})$$

or
$$y = \frac{N}{(n-3)(n-5)\dots 4.2} \left(\frac{n}{\sigma^2}\right)^{\frac{n-1}{2}} x^{n-2} e^{-\frac{nx^2}{2\sigma^2}} \quad (n \text{ odd})$$

where $N$ as usual represents the total frequency.

## Section II.

To show that there is no correlation between (*a*) the distance of the mean of a sample from the mean of the population and (*b*) the standard deviation of a sample with normal distribution.

(1) Clearly positive and negative positions of the mean of the sample are equally likely, and hence there cannot be correlation between the absolute value of the distance of the mean from the mean of the population and the standard deviation, but (2) there might be correlation between the square of the distance and the square of the standard deviation.

. Let
$$u^2 = \left(\frac{S(x_1)}{n}\right)^2 \quad \text{and} \quad s^2 = \frac{S(x_1^2)}{n} - \left(\frac{S(x_1)}{n}\right)^2.$$

Then if $m_1'$, $M_1'$ be the mean values of $u^2$ and $s^2$, we have by the preceding part $M_1' = \mu_2 \frac{(n-1)}{n}$ and $m_1' = \frac{\mu_2}{n}$.

$$\text{Now } u^2s^2 = \frac{S(x_1^2)}{n}\left(\frac{S(x_1)}{n}\right)^2 - \left(\frac{S(x_1)}{n}\right)^4$$

$$= \left(\frac{S(x_1^2)}{n}\right)^2 + 2\,\frac{S(x_1x_2).S(x_1^2)}{n^3} - \frac{S(x_1^4)}{n^4}$$

$$- \frac{6x_1^2x_2^2}{n^4} - \text{other terms of odd order which will vanish on summation.}$$

Summing for all values and dividing by the number of cases we get

$$R_{u^2s^2}\sigma_{u^2}\sigma_{s^2} + m_1M_1 = \frac{\mu_4}{n^2} + \mu_2^2\frac{(n-1)}{n^3} - \frac{\mu_4}{n^3} - 3\mu_2^2\frac{(n-1)}{n^3},$$

where $R_{u^2s^2}$ is the correlation between $u^2$ and $s^2$.

$$R_{u^2s^2}\sigma_{u^2}\sigma_{s^2} + \mu_2^2\frac{(n-1)}{n^3} = \mu_2^2\frac{(n-1)}{n^3}\{3+n-3\} = \mu_2^2\frac{(n-1)}{n^2}.$$

Hence $R_{u^2s^2}\sigma_{u^2}\sigma_{s^2} = 0$ or there is no correlation between $u^2$ and $s^2$.

## SECTION III.

To find the equation representing the frequency distribution of the means of samples of $n$ drawn from a normal population, the mean being expressed in terms of the standard deviation of the sample.

We have $y = \dfrac{C}{\sigma^{n-1}}\,s^{n-2}e^{-\frac{ns^2}{2\sigma^2}}$ as the equation representing the distribution of $s$, the standard deviation of a sample of $n$, when the samples are drawn from a normal population with standard deviation $\sigma$.

Now the means of these samples of $n$ are distributed according to the equation

$$y = \frac{\sqrt{n}N}{\sqrt{2\pi}\,\sigma}\,e^{-\frac{nx^2}{2\sigma^2}}*$$

and we have shown that there is no correlation between $x$, the distance of the mean of the sample, and $s$, the standard deviation of the sample.

Now let us suppose $x$ measured in terms of $s$, *i.e.* let us find the distribution of $z = \dfrac{x}{s}$.

If we have $y_1 = \phi(x)$ and $y_2 = \psi(z)$ as the equations representing the frequency of $x$ and of $z$ respectively, then

$$y_1dx = y_2dz = y_2\frac{dx}{s},$$

$$\therefore\ y_2 = sy_1.$$

* Airy, *Theory of Errors of Observations*, Part II. § 6.

Hence $$y = \frac{N\sqrt{n}s}{\sqrt{2\pi}\sigma} e^{-\frac{ns^2z^2}{2\sigma^2}}$$

is the equation representing the distribution of $z$ for samples of $n$ with standard deviation $s$.

Now the chance that $s$ lies between $s$ and $s+ds$ is:

$$\frac{\int_s^{s+ds} \frac{C}{\sigma^{n-1}} s^{n-2} e^{-\frac{ns^2}{2\sigma^2}} ds}{\int_0^{\infty} \frac{C}{\sigma^{n-1}} s^{n-2} e^{-\frac{ns^2}{2\sigma^2}} ds},$$

which represents the $N$ in the above equation.

Hence the distribution of $z$ due to values of $s$ which lie between $s$ and $s+ds$ is

$$y = \frac{\int_s^{s+ds} \frac{C}{\sigma^n}\sqrt{\frac{n}{2\pi}} s^{n-1} e^{-\frac{ns^2(1+z^2)}{2\sigma^2}} ds}{\int_0^{\infty} \frac{C}{\sigma^{n-1}} s^{n-2} e^{-\frac{ns^2}{2\sigma^2}} ds} = \frac{\sqrt{\frac{n}{2\pi}} \int_s^{s+ds} s^{n-1} e^{-\frac{ns^2(1+z^2)}{2\sigma^2}} ds}{\sigma \int_0^{\infty} s^{n-2} e^{-\frac{ns^2}{2\sigma^2}} ds},$$

and summing for all values of $s$ we have as an equation giving the distribution of $z$

$$y = \frac{\sqrt{\frac{n}{2\pi}}}{\sigma} \frac{\int_0^{\infty} s^{n-1} e^{-\frac{ns^2(1+z^2)}{2\sigma^2}} ds}{\int_0^{\infty} s^{n-2} e^{-\frac{ns^2}{2\sigma^2}} ds}.$$

By what we have already proved this reduces to

$$y = \frac{1}{2}\frac{n-2}{n-3}\cdot\frac{n-4}{n-5}\cdots\frac{5}{4}\cdot\frac{3}{2}(1+z^2)^{-\frac{n}{2}} \text{ if } n \text{ be odd,}$$

and to $$y = \frac{1}{\pi}\frac{n-2}{n-3}\cdot\frac{n-4}{n-5}\cdots\frac{4}{3}\cdot\frac{2}{1}(1+z^2)^{-\frac{n}{2}} \text{ if } n \text{ be even.}$$

Since this equation is independent of $\sigma$ it will give the distribution of the distance of the mean of a sample from the mean of the population expressed in terms of the standard deviation of the sample for any normal population.

## Section IV.

### *Some Properties of the Standard Deviation Frequency Curve.*

By a similar method to that adopted for finding the constant we may find the mean and moments: thus the mean is at $\frac{I_{n-1}}{I_{n-2}}$,

which is equal to $$\frac{(n-2)(n-4)}{(n-3)(n-5)}\cdots\frac{2}{1}\sqrt{\frac{2}{\pi}}\frac{\sigma}{\sqrt{n}} \text{ (if } n \text{ be even),}$$

or $$\frac{(n-2)(n-4)}{(n-3)(n-5)}\cdots\frac{3}{2}\sqrt{\frac{\pi}{2}}\frac{\sigma}{\sqrt{n}} \text{ (if } n \text{ be odd).}$$

The second moment about the end of the range is

$$\frac{I_n}{I_{n-2}} = \frac{(n-1)\,\sigma^2}{n}.$$

The third moment about the end of the range is equal to

$$\frac{I_{n+1}}{I_{n-2}} = \frac{I_{n+1}}{I_{n-1}} \cdot \frac{I_{n-1}}{I_{n-2}}$$
$$= \sigma^2 \times \text{the mean.}$$

The fourth moment about the end of the range is equal to

$$\frac{I_{n+2}}{I_{n-2}} = \frac{(n-1)(n+1)}{n^2}\,\sigma^4.$$

If we write the distance of the mean from the end of the range $\frac{D\sigma}{\sqrt{n}}$ and the moments about the end of the range $\nu_1$, $\nu_2$, etc.

then $$\nu_1 = \frac{D\sigma}{\sqrt{n}}, \quad \nu_2 = \frac{n-1}{n}\,\sigma^2, \quad \nu_3 = \frac{D\sigma^3}{\sqrt{n}}, \quad \nu_4 = \frac{n^2-1}{n^2}\,\sigma^4.$$

From this we get the moments about the mean

$$\mu_2 = \frac{\sigma^2}{n}(n - 1 - D^2),$$

$$\mu_3 = \frac{\sigma^3}{n\sqrt{n}}\{nD - 3(n-1)D + 2D^3\} = \frac{\sigma^3 D}{n\sqrt{n}}\{2D^2 - 2n + 3\},$$

$$\mu_4 = \frac{\sigma^4}{n^2}\{n^2 - 1 - 4D^2 n + 6(n-1)D^2 - 3D^4\} = \frac{\sigma^4}{n^2}\{n^2 - 1 - D^2(3D^2 - 2n + 6)\}.$$

It is of interest to find out what these become when $n$ is large.

In order to do this we must find out what is the value of $D$.

Now Wallis's expression for $\pi$ derived from the infinite product value of $\sin x$ is

$$\frac{\pi}{2}(2n+1) = \frac{2^2 \cdot 4^2 \cdot 6^2 \ldots (2n)^2}{1^2 \cdot 3^2 \cdot 5^2 \ldots (2n-1)^2}.$$

If we assume a quantity $\theta\left(= a_0 + \frac{a_1}{n} + \text{etc.}\right)$ which we may add to the $2n+1$ in order to make the expression approximate more rapidly to the truth, it is easy to show that $\theta = -\frac{1}{2} + \frac{1}{16n} - \text{etc.}$ and we get

$$\frac{\pi}{2}\left(2n + \frac{1}{2} + \frac{1}{16n}\right) = \frac{2^2 \cdot 4^2 \cdot 6^2 \ldots (2n)^2}{1^2 \cdot 3^2 \cdot 5^2 \ldots (2n-1)^2}.$$ *

From this we find that whether $n$ be even or odd $D^2$ approximates to $n - \frac{3}{2} + \frac{1}{8n}$ when $n$ is large.

* This expression will be found to give a much closer approximation to $\pi$ than Wallis's.

Substituting this value of $D$ we get

$$\mu_2 = \frac{\sigma^2}{2n}\left(1 - \frac{1}{4n}\right), \qquad \mu_3 = \frac{\sigma^3\sqrt{1 - \frac{3}{2n} + \frac{1}{16n^2}}}{4n^2}, \qquad \mu_4 = \frac{3\sigma^4}{4n^2}\left(1 + \frac{1}{2n} - \frac{1}{16n^2}\right).$$

Consequently the value of the standard deviation of a standard deviation which we have found $\left(\dfrac{\sigma}{\sqrt{2n}\sqrt{1 - \frac{1}{4n}}}\right)$ becomes the same as that found for the normal curve by Professor Pearson $(\sigma/\sqrt{2n})$ when $n$ is large enough to neglect the $1/4n$ in comparison with 1.

Neglecting terms of lower order than $\dfrac{1}{n}$ we find

$$\beta_1 = \frac{2n-3}{n(4n-3)}, \qquad \beta_2 = 3\left(1 - \frac{1}{2n}\right)\left(1 + \frac{1}{2n}\right).$$

Consequently as $n$ increases $\beta_2$ very soon approaches the value 3 of the normal curve, but $\beta_1$ vanishes more slowly, so that the curve remains slightly skew.

DIAGRAM I. Frequency curve giving the distribution of Standard Deviations of samples of 10 taken from a normal population.

Equation $y = \dfrac{N}{7.5.3}\dfrac{10^{\frac{9}{2}}}{\sigma^9}\sqrt{\dfrac{2}{\pi}}\,x^8 e^{-\frac{10x^2}{2\sigma^2}}$.

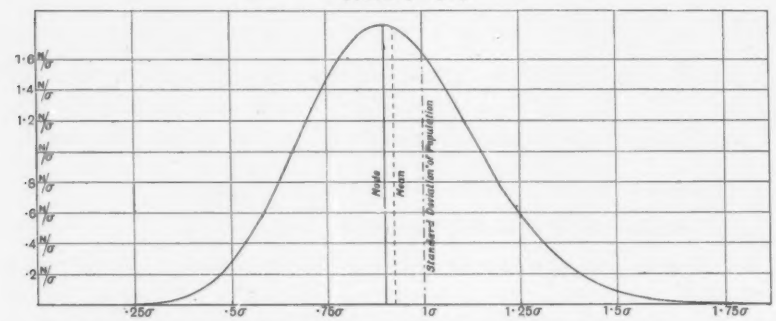



Diagram I shows the theoretical distribution of the S.D. found from samples of 10.

$$y = \frac{N10^{\frac{9}{2}}}{7.5.3}\sqrt{\frac{2}{\pi}}\frac{x^8}{\sigma^9}e^{-\frac{10x^2}{2\sigma^2}}.$$

## SECTION V.

*Some properties of the curve* $y = \dfrac{n-2}{n-3}\cdot\dfrac{n-4}{n-5}\cdots\begin{pmatrix}\frac{4}{3}\cdot\frac{2}{\pi} \text{ if } n \text{ be even}\\ \frac{5}{4}\cdot\frac{3}{2} \text{ if } n \text{ be odd}\end{pmatrix}(1+z^2)^{-\frac{n}{2}}$.

Writing $z = \tan\theta$ the equation becomes $y = \dfrac{n-2}{n-3}\cdot\dfrac{n-4}{n-5}\cdots$ etc. $\times \cos^n\theta$, which affords an easy way of drawing the curve. Also $dz = d\theta/\cos^2\theta$.

Hence to find the area of the curve between any limits we must find

$$\frac{n-2}{n-3}\cdot\frac{n-4}{n-5}\ldots \text{etc.}\times\int\cos^{n-2}\theta\,d\theta$$

$$=\frac{n-2}{n-3}\cdot\frac{n-4}{n-5}\ldots \text{etc.}\left\{\frac{n-3}{n-2}\int\cos^{n-4}\theta\,d\theta+\left[\frac{\cos^{n-3}\theta\sin\theta}{n-2}\right]\right\}$$

$$=\frac{n-4}{n-5}\cdot\frac{n-6}{n-7}\ldots \text{etc.}\int\cos^{n-4}\theta\,d\theta+\frac{1}{n-3}\cdot\frac{n-4}{n-5}\ldots \text{etc.}\,[\cos^{n-3}\theta\sin\theta],$$

and by continuing the process the integral may be evaluated.

For example, if we wish to find the area between 0 and $\theta$ for $n=8$ we have

$$\text{area}=\frac{6}{5}\cdot\frac{4}{3}\cdot\frac{2}{1}\cdot\frac{1}{\pi}\int_0^\theta\cos^6\theta\,d\theta$$

$$=\frac{4}{3}\cdot\frac{2}{\pi}\cdot\int_0^\theta\cos^4\theta\,d\theta+\frac{1}{5}\cdot\frac{4}{3}\cdot\frac{2}{\pi}\cos^5\theta\sin\theta$$

$$=\frac{\theta}{\pi}+\frac{1}{\pi}\cos\theta\sin\theta+\frac{1}{3}\cdot\frac{2}{\pi}\cos^3\theta\sin\theta+\frac{1}{5}\cdot\frac{4}{3}\cdot\frac{2}{\pi}\cos^5\theta\sin\theta,$$

and it will be noticed that for $n=10$ we shall merely have to add to this same expression the term $\frac{1}{7}\cdot\frac{6}{5}\cdot\frac{4}{3}\cdot\frac{2}{\pi}\cos^7\theta\sin\theta$.

The tables at the end of the paper give the area between $-\infty$ and $z$

$$\left(\text{or }\theta=-\frac{\pi}{2}\text{ and }\theta=\tan^{-1}z\right).$$

This is the same as $\cdot5+$ the area between $\theta=0$, and $\theta=\tan^{-1}z$, and as the whole area of the curve is equal to 1, the tables give the probability that the mean of the sample does not differ by more than $z$ times the standard deviation of the sample from the mean of the population.

The whole area of the curve is equal to

$$\frac{n-2}{n-3}\cdot\frac{n-4}{n-5}\ldots \text{etc.}\times\int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}}\cos^{n-2}\theta\,d\theta,$$

and since all the parts between the limits vanish at both limits this reduces to 1.

Similarly the second moment coefficient is equal to

$$\frac{n-2}{n-3}\cdot\frac{n-4}{n-5}\ldots \text{etc.}\times\int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}}\cos^{n-2}\theta\tan^2\theta\,d\theta$$

$$=\frac{n-2}{n-3}\cdot\frac{n-4}{n-5}\ldots \text{etc.}\times\int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}}(\cos^{n-4}\theta-\cos^{n-2}\theta)\,d\theta$$

$$=\frac{n-2}{n-3}-1=\frac{1}{n-3}.$$

Hence the standard deviation of the curve is $1/\sqrt{n-3}$. The fourth moment coefficient is equal to

$$\frac{n-2}{n-3}\cdot\frac{n-4}{n-5}\ldots \text{etc.} \times \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} \cos^{n-2}\theta \tan^4\theta\, d\theta$$

$$= \frac{n-2}{n-3}\cdot\frac{n-4}{n-5}\ldots \text{etc.} \times \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} (\cos^{n-6}\theta - 2\cos^{n-4}\theta + \cos^{n-2}\theta)\, d\theta$$

$$= \frac{n-2}{n-3}\cdot\frac{n-4}{n-5} - \frac{2(n-2)}{n-3} + 1 = \frac{3}{(n-3)(n-5)}.$$

The odd moments are of course zero as the curve is symmetrical, so

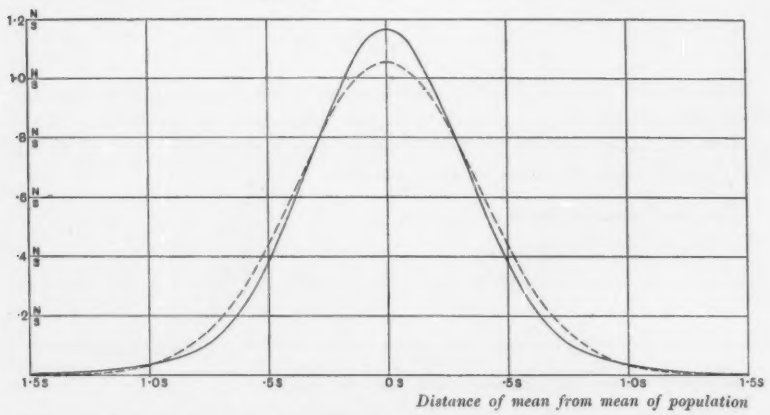
$$\beta_1 = 0, \qquad \beta_2 = \frac{3(n-3)}{n-5} = 3 + \frac{2}{n-5}.$$

Hence as $n$ increases the curve approaches the normal curve whose standard deviation is $1/\sqrt{n-3}$.

$\beta_2$ however is always greater than 3, indicating that large deviations are more common than in the normal curve.

DIAGRAM II. Solid curve $y = \frac{N}{S} \times \frac{8}{7}\cdot\frac{6}{5}\cdot\frac{4}{3}\cdot\frac{2}{\pi}\cos^{10}\theta, \quad x/s = \tan\theta.$

Broken line curve $y = \frac{\sqrt{7}\,.\,N}{\sqrt{2\pi}\,.\,s}\, e^{-\frac{7x^2}{2s^2}}$, the normal curve with the same S.D.




I have tabled the area for the normal curve with standard deviation $1/\sqrt{7}$ so as to compare with my curve for $n = 10$*. It will be seen that odds laid according to either table would not seriously differ till we reach $z = \cdot 8$, where the odds are about 50 to 1 that the mean is within that limit: beyond that the normal curve gives a false feeling of security, for example, according to the normal curve it is 99,986 to 14 (say 7000 to 1) that the mean of the population lies between $-\infty$ and $+1\cdot 3s$ whereas the real odds are only 99,819 to 181 (about 550 to 1).

* See p. 19.

Now 50 to 1 corresponds to three times the probable error in the normal curve and for most purposes would be considered significant; for this reason I have only tabled my curves for values of $n$ not greater than 10, but have given the $n=9$ and $n=10$ tables to one further place of decimals. They can be used as foundations for finding values for larger samples*.

The table for $n=2$ can be readily constructed by looking out $\theta=\tan^{-1}z$ in Chambers' Tables and then $\cdot5+\theta/\pi$ gives the corresponding value.

Similarly $\frac{1}{2}\sin\theta+\cdot5$ gives the values when $n=3$.

There are two points of interest in the $n=2$ curve. Here $s$ is equal to half the distance between the two observations. $\tan^{-1}\frac{s}{s}=\frac{\pi}{4}$ so that between $+s$ and $-s$ lies $2\times\frac{\pi}{4}\times\frac{1}{\pi}$ or half the probability, i.e. if two observations have been made and we have no other information, it is an even chance that the mean of the (normal) population will lie between them. On the other hand the second moment coefficient is

$$\frac{1}{\pi}\int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}}\tan^2\theta\,d\theta=\frac{1}{\pi}\left[\tan\theta-\theta\right]_{-\frac{\pi}{2}}^{\frac{\pi}{2}}=\infty,$$

or the standard deviation is infinite while the probable error is finite.

SECTION VI. *Practical Test of the foregoing Equations.*

Before I had succeeded in solving my problem analytically, I had endeavoured to do so empirically. The material used was a correlation table containing the height and left middle finger measurements of 3000 criminals, from a paper by W. R. Macdonell (*Biometrika*, Vol. I. p. 219). The measurements were written out on 3000 pieces of cardboard, which were then very thoroughly shuffled and drawn at random. As each card was drawn its numbers were written down in a book which thus contains the measurements of 3000 criminals in a random order. Finally each consecutive set of 4 was taken as a sample—750 in all—and the mean, standard deviation, and correlation† of each sample determined. The difference between the mean of each sample and the mean of the population was then divided by the standard deviation of the sample, giving us the $z$ of Section III.

This provides us with two sets of 750 standard deviations and two sets of 750 $z$'s on which to test the theoretical results arrived at. The height and left middle finger correlation table was chosen because the distribution of both was approximately normal and the correlation was fairly high. Both frequency curves, however, deviate slightly from normality, the constants being for height $\beta_1=\cdot0026$, $\beta_2=3\cdot175$, and for left middle finger lengths $\beta_1=\cdot0030$, $\beta_2=3\cdot140$, and in consequence there is a tendency for a certain number of larger standard deviations to occur than if the distributions were normal. This, however, appears to make very little difference to the distribution of $z$.

* E.g. if $n=11$, to the corresponding value for $n=9$, we add $\frac{7}{8}\times\frac{8}{9}\times\frac{9}{10}\times\frac{1}{2}\times\frac{1}{2}\cos^8\theta\sin\theta$; if $n=13$ we add as well $\frac{9}{10}\times\frac{7}{8}\times\frac{8}{9}\times\frac{9}{10}\times\frac{1}{2}\times\frac{1}{2}\cos^{10}\theta\sin\theta$ and so on.

† I hope to publish the results of the correlation work shortly.

Another thing which interferes with the comparison is the comparatively large groups in which the observations occur. The heights are arranged in 1 inch groups, the standard deviation being only 2·54 inches: while the finger lengths were originally grouped in millimetres, but unfortunately I did not at the time see the importance of having a smaller unit, and condensed them into two millimetre groups, in terms of which the standard deviation is 2·74.

Several curious results follow from taking samples of 4 from material disposed in such wide groups. The following points may be noticed:

(1) The means only occur as multiples of ·25.

(2) The standard deviations occur as the square roots of the following types of numbers $n$, $n + ·19$, $n + ·25$, $n + ·50$, $n + ·69$, $2n + ·75$.

(3) A standard deviation belonging to one of these groups can only be associated with a mean of a particular kind; thus a standard deviation of $\sqrt{2}$ can only occur if the mean differs by a whole number from the group we take as origin, while $\sqrt{1·69}$ will only occur when the mean is at $n \pm ·25$.

(4) All the four individuals of the sample will occasionally come from the same group, giving a zero value for the standard deviation. Now this leads to an infinite value of $z$ and is clearly due to too wide a grouping, for although two men may have the same height when measured by inches, yet the finer the measurements the more seldom will they be identical, till finally the chance that four men will have *exactly* the same height is infinitely small. If we had smaller grouping the zero values of the standard deviation might be expected to increase, and a similar consideration will show that the smaller values of the standard deviation would also be likely to increase, such as ·436, when 3 fall in one group and 1 in an adjacent group, or ·50 when 2 fall in two adjacent groups. On the other hand when the individuals of the sample lie far apart, the argument of Sheppard's correction will apply, the real value of the standard deviation being more likely to be smaller than that found owing to the frequency in any group being greater on the side nearer the mode.

These two effects of grouping will tend to neutralise each other in their effect on the mean value of the standard deviation, but both will increase the variability.

Accordingly we find that the mean value of the standard deviation is quite close to that calculated, while in each case the variability is sensibly greater. The fit of the curve is not good, both for this reason and because the frequency is not evenly distributed owing to effects (2) and (3) of grouping. On the other hand the fit of the curve giving the frequency of $z$ is very good and as that is the only practical point the comparison may be considered satisfactory.

The following are the figures for height:—

| | | |
|---|---|---|
| Mean value of standard deviations; | calculated | 2·027 ± ·021 |
| „ „ „ | observed | 2·026 |
| | Difference = | – ·001 |

Standard deviation of standard deviations:—

Calculated ·8556 ± ·015
Observed ·9066
Difference = + ·0510

*Comparison of Fit. Theoretical Equation:* $y = \frac{16 \times 750}{\sqrt{2\pi\sigma^3}} x^2 e^{-\frac{2x^2}{\sigma^2}}$.

| Scale in terms of standard deviation of population | 0 to ·1 | ·1 to ·2 | ·2 to ·3 | ·3 to ·4 | ·4 to ·5 | ·5 to ·6 | ·6 to ·7 | ·7 to ·8 | ·8 to ·9 | ·9 to 1·0 | 1·0 to 1·1 | 1·1 to 1·2 | 1·2 to 1·3 | 1·3 to 1·4 | 1·4 to 1·5 | 1·5 to 1·6 | 1·6 to 1·7 | Greater than 1·7 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Calculated frequency | 1½ | 10½ | 27 | 45½ | 64½ | 78½ | 87 | 88 | 81½ | 71 | 58 | 45 | 33 | 23 | 15 | 9½ | 5½ | 7 |
| Observed frequency | 3 | 14½ | 24½ | 37½ | 107 | 67 | 73 | 77 | 77½ | 64 | 52½ | 49½ | 35 | 28 | 12½ | 9 | 11½ | 7 |
| Difference | +1½ | +4 | −2½ | −8 | +42½ | −11½ | −14 | −11 | −4 | −7 | −5½ | +4½ | +2 | +5 | −2½ | −½ | +6 | 0 |

whence $\chi^2 = 48{\cdot}06$, $P = {\cdot}000{,}06$ (about).

In tabling the observed frequency, values between ·0125 and ·0875 were included in one group, while between ·0875 and ·0125 they were divided over the two groups. As an instance of the irregularity due to grouping I may mention that there were 31 cases of standard deviations 1·30 (in terms of the grouping) which is ·5117 in terms of the standard deviation of the population, and they were therefore divided over the groups ·4 to ·5 and ·5 to ·6. Had they all been counted in groups ·5 to ·6 $\chi^2$ would have fallen to 29·85 and $P$ would have risen to ·03. The $\chi^2$ test presupposes *random* sampling from a frequency following the given law, but this we have not got owing to the interference of the grouping.

When, however, we test the $z$'s where the grouping has not had so much effect we find a close correspondence between the theory and the actual result.

There were three cases of infinite values of $z$ which, for the reasons given above, were given the next largest values which occurred, namely + 6 or − 6. The rest were divided into groups of ·1; ·04, ·05 and ·06, being divided between the two groups on either side.

The calculated value for the standard deviation of the frequency curve was 1 (± ·017) while the observed was 1·039. The value of the standard deviation is really infinite, as the fourth moment coefficient is infinite, but as we have arbitrarily limited the infinite cases we may take as an approximation $\frac{1}{\sqrt{1500}}$ from which the value of the probable error given above is obtained. The fit of the curve is as follows:—

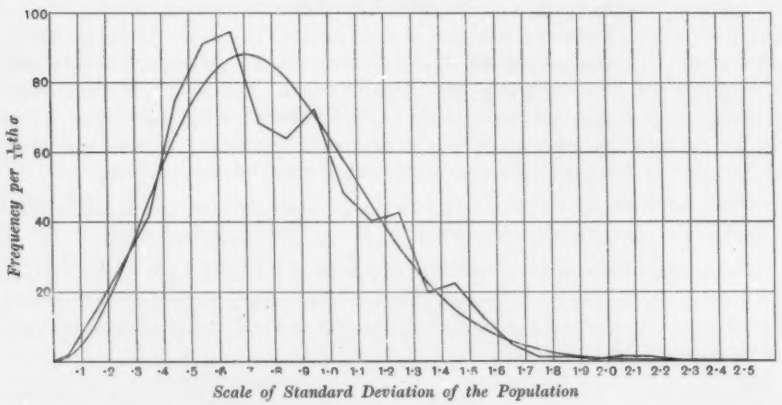
*Comparison of Fit. Theoretical Equation:* $y = \frac{2}{\pi}\cos^4\theta$, $z = \tan\theta$.

| Scale of $z$ | less than −3·05 | −3·05 to −2·05 | −2·05 to −1·55 | −1·55 to −1·05 | −1·05 to −·75 | −·75 to −·45 | −·45 to −·15 | −·15 to +·15 | +·15 to +·45 | +·45 to +·75 | +·75 to +1·05 | +1·05 to +1·55 | +1·55 to +2·05 | +2·05 to +3·05 | more than +3·05 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Calculated frequency | 5 | 9½ | 13½ | 34½ | 44½ | 78½ | 119 | 141 | 119 | 78½ | 44½ | 34½ | 13½ | 9½ | 5 |
| Observed frequency | 9 | 14½ | 11½ | 33 | 43½ | 70½ | 119½ | 151½ | 122 | 67½ | 49 | 26½ | 16 | 10 | 6 |
| Difference | +4 | +5 | −2 | −1½ | −1 | −8 | +½ | +10½ | +3 | −11 | +4½ | −8 | +2½ | +½ | +1 |

whence $\chi^2 = 12{\cdot}44$, $P = {\cdot}56$.

This is very satisfactory, especially when we consider that as a rule observations are tested against curves fitted from the mean and one or more other moments of the observations, so that considerable correspondence is only to be expected; while this curve is exposed to the full errors of random sampling, its constants having been calculated quite apart from the observations.

Diagram III. Comparison of Calculated Standard Deviation Frequency Curve with 750 actual Standard Deviations.




The left middle finger samples show much the same features as those of the height, but as the grouping is not so large compared to the variability the curves fit the observations more closely. Diagrams III.* and IV. give the standard deviations and the $z$'s for this set of samples. The results are as follows:—

* There are three small mistakes in plotting the observed values in Diagram III., which make the fit appear worse than it really is.

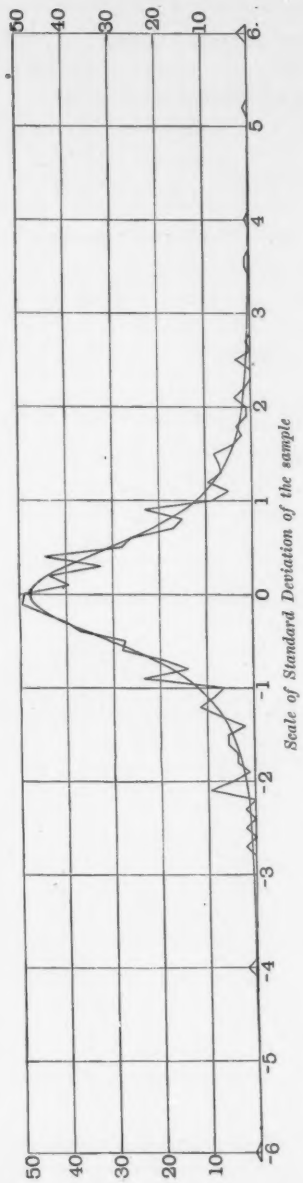



Diagram IV. Comparison of the theoretical frequency curve $y = \frac{1500}{\pi}\left(1+\frac{x^2}{s^2}\right)^{-2}$, with an actual sample of 750 cases.

Mean value of standard deviations; calculated $2{\cdot}186 \pm {\cdot}023$
" " " observed $2{\cdot}179$
Difference $= -\ {\cdot}007$

Standard deviation of standard deviations:—
Calculated ${\cdot}9224 \pm {\cdot}016$
Observed ${\cdot}9802$
Difference $= +\ {\cdot}0578$

*Comparison of Fit. Theoretical Equation:* $y = \dfrac{16 \times 750}{\sqrt{2\pi}\sigma^3} x^2 e^{-\frac{2x^2}{\sigma^2}}$.

| Scale in terms of standard deviation of population | 0 to ·1 | ·1 to ·2 | ·2 to ·3 | ·3 to ·4 | ·4 to ·5 | ·5 to ·6 | ·6 to ·7 | ·7 to ·8 | ·8 to ·9 | ·9 to 1·0 | 1·0 to 1·1 | 1·1 to 1·2 | 1·2 to 1·3 | 1·3 to 1·4 | 1·4 to 1·5 | 1·5 to 1·6 | 1·6 to 1·7 | greater than 1·7 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Calculated frequency | 1½ | 10½ | 27 | 45½ | 64½ | 78½ | 87 | 88 | 81½ | 71 | 58 | 45 | 33 | 23 | 15 | 9½ | 5½ | 7 |
| Observed frequency | 2 | 14 | 27½ | 51 | 64½ | 91 | 94½ | 68½ | 65½ | 73 | 48½ | 40½ | 42½ | 20 | 22½ | 12 | 5 | 7½ |
| Difference | +½ | +3½ | +½ | +5½ | — | +12½ | +7½ | −19½ | −16 | +2 | −9½ | −4½ | +9½ | −3 | +7½ | +2½ | −½ | +½ |

whence $\chi^2 = 21{\cdot}80$, $P = {\cdot}19$.

Calculated value of standard deviation $1\ (\pm {\cdot}017)$
Observed " " " ${\cdot}982$
Difference $= -{\cdot}018$

*Comparison of Fit. Theoretical Equation:* $y = \dfrac{2}{\pi}\cos^4\theta, \quad z = \tan\theta.$

| Scale of $z$ | less than −3·05 | −3·05 to −2·05 | −2·05 to −1·55 | −1·55 to −1·05 | −1·05 to −·75 | −·75 to −·45 | −·45 to −·15 | −·15 to +·15 | +·15 to +·45 | +·45 to +·75 | +·75 to +1·05 | +1·05 to +1·55 | +1·55 to +2·05 | +2·05 to +3·05 | more than +3·05 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Calculated frequency | 5 | 9½ | 13½ | 34½ | 44½ | 78½ | 119 | 141 | 119 | 78½ | 44½ | 34½ | 13½ | 9½ | 5 |
| Observed frequency | 4 | 15½ | 18 | 33½ | 44 | 75 | 122 | 138 | 120½ | 71 | 46½ | 36 | 11 | 9 | 6 |
| Difference | −1 | +6 | +4½ | −1 | −½ | −3½ | +3 | −3 | +1½ | −7½ | +2 | +1½ | −2½ | −½ | +1 |

whence $\chi^2 = 7{\cdot}39$, $P = {\cdot}92$.

A very close fit.

We see then that if the distribution is approximately normal our theory gives us a satisfactory measure of the certainty to be derived from a small sample in both the cases we have tested; but we have an indication that a fine grouping is

of advantage. If the distribution is not normal, the mean and the standard deviation of a sample will be positively correlated, so that although both will have greater variability, yet they will tend to counteract each other, a mean deviating largely from the general mean tending to be divided by a larger standard deviation Consequently I believe that the tables at the end of the present paper may be used in estimating the degree of certainty arrived at by the mean of a few experiments, in the case of most laboratory or biological work where the distributions are as a rule of a 'cocked hat' type and so sufficiently nearly normal.

SECTION VII. *Tables of* $\dfrac{n-2}{n-3}\dfrac{n-4}{n-5}\cdots\begin{pmatrix}\frac{3}{2}\cdot\frac{1}{2}n \text{ odd}\\ \frac{2}{1}\cdot\frac{1}{\pi}n \text{ even}\end{pmatrix}\int_{-\frac{\pi}{2}}^{\tan^{-1}x}\cos^{n-2}\theta\,d\theta$

*for values of* $n$ *from* 4 *to* 10 *inclusive.*

*Together with* $\dfrac{\sqrt{7}}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{7x^2}{2}}dx$ *for comparison when* $n=10$.

| $z\left(=\frac{x}{s}\right)$ | $n=4$ | $n=5$ | $n=6$ | $n=7$ | $n=8$ | $n=9$ | $n=10$ | For comparison $\left(\frac{\sqrt{7}}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{7x^2}{2}}dx\right)$ |
|---|---|---|---|---|---|---|---|---|
| ·1 | ·5633 | ·5745 | ·5841 | ·5928 | ·6006 | ·60787 | ·61462 | ·60411 |
| ·2 | ·6241 | ·6458 | ·6634 | ·6798 | ·6936 | ·70705 | ·71846 | ·70159 |
| ·3 | ·6804 | ·7096 | ·7340 | ·7549 | ·7733 | ·78961 | ·80423 | ·78641 |
| ·4 | ·7309 | ·7657 | ·7939 | ·8175 | ·8376 | ·85465 | ·86970 | ·85520 |
| ·5 | ·7749 | ·8131 | ·8428 | ·8667 | ·8863 | ·90251 | ·91609 | ·90691 |
| ·6 | ·8125 | ·8518 | ·8813 | ·9040 | ·9218 | ·93600 | ·94732 | ·94375 |
| ·7 | ·8440 | ·8830 | ·9109 | ·9314 | ·9468 | ·95851 | ·96747 | ·96799 |
| ·8 | ·8701 | ·9076 | ·9332 | ·9512 | ·9640 | ·97328 | ·98007 | ·98253 |
| ·9 | ·8915 | ·9269 | ·9498 | ·9652 | ·9756 | ·98279 | ·98780 | ·99137 |
| 1·0 | ·9092 | ·9419 | ·9622 | ·9751 | ·9834 | ·98890 | ·99252 | ·99820 |
| 1·1 | ·9236 | ·9537 | ·9714 | ·9821 | ·9887 | ·99280 | ·99539 | ·99926 |
| 1·2 | ·9354 | ·9628 | ·9782 | ·9870 | ·9922 | ·99528 | ·99713 | ·99971 |
| 1·3 | ·9451 | ·9700 | ·9832 | ·9905 | ·9946 | ·99688 | ·99819 | ·99986 |
| 1·4 | ·9531 | ·9756 | ·9870 | ·9930 | ·9962 | ·99791 | ·99885 | ·99989 |
| 1·5 | ·9598 | ·9800 | ·9899 | ·9948 | ·9973 | ·99859 | ·99926 | ·99999 |
| 1·6 | ·9653 | ·9836 | ·9920 | ·9961 | ·9981 | ·99903 | ·99951 | |
| 1·7 | ·9699 | ·9864 | ·9937 | ·9970 | ·9986 | ·99933 | ·99968 | |
| 1·8 | ·9737 | ·9886 | ·9950 | ·9977 | ·9990 | ·99953 | ·99978 | |
| 1·9 | ·9770 | ·9904 | ·9959 | ·9983 | ·9992 | ·99967 | ·99985 | |
| 2·0 | ·9797 | ·9919 | ·9967 | ·9986 | ·9994 | ·99976 | ·99990 | |
| 2·1 | ·9821 | ·9931 | ·9973 | ·9989 | ·9996 | ·99983 | ·99993 | |
| 2·2 | ·9841 | ·9941 | ·9978 | ·9992 | ·9997 | ·99987 | ·99995 | |
| 2·3 | ·9858 | ·9950 | ·9982 | ·9993 | ·9998 | ·99991 | ·99996 | |
| 2·4 | ·9873 | ·9957 | ·9985 | ·9995 | ·9998 | ·99993 | ·99997 | |
| 2·5 | ·9886 | ·9963 | ·9987 | ·9996 | ·9998 | ·99995 | ·99998 | |
| 2·6 | ·9898 | ·9967 | ·9989 | ·9996 | ·9999 | ·99996 | ·99999 | |
| 2·7 | ·9908 | ·9972 | ·9991 | ·9997 | ·9999 | ·99997 | ·99999 | |
| 2·8 | ·9916 | ·9975 | ·9992 | ·9998 | ·9999 | ·99998 | ·99999 | |
| 2·9 | ·9924 | ·9978 | ·9993 | ·9998 | ·9999 | ·99998 | ·99999 | |
| 3·0 | ·9931 | ·9981 | ·9994 | ·9998 | — | ·99999 | — | — |

### Section VIII. *Explanation of Tables.*

The tables give the probability that the value of the mean, measured from the mean of the population, in terms of the standard deviation of the sample, will lie between $-\infty$ and $z$. Thus, to take the table for samples of six, the probability of the mean of the population lying between $-\infty$ and once the standard deviation of the sample is ·9622 or the odds are about 24 to 1 that the mean of the population lies between these limits.

The probability is therefore ·0378 that it is greater than once the standard deviation and ·0756 that it lies outside $\pm 1{\cdot}0$ times the standard deviation.

### Section IX. *Illustrations of Method.*

*Illustration I.* As an instance of the kind of use which may be made of the tables, I take the following figures from a table by A. R. Cushny and A. R. Peebles in the *Journal of Physiology* for 1904, showing the different effects of the optical isomers of hyoscyamine hydrobromide in producing sleep. The sleep of 10 patients was measured without hypnotic and after treatment (1) with D. hyoscyamine hydrobromide, (2) with L. hyoscyamine hydrobromide. The average number of hours' sleep gained by the use of the drug is tabulated below.

The conclusion arrived at was that in the usual dose 2 was, but 1 was not, of value as a soporific.

*Additional hours' sleep gained by the use of hyoscyamine hydrobromide.*

| Patient | 1 (Dextro-) | 2 (Laevo-) | Difference (2–1) |
|---|---|---|---|
| 1. | + ·7 | + 1·9 | + 1·2 |
| 2. | − 1·6 | + ·8 | + 2·4 |
| 3. | − ·2 | + 1·1 | + 1·3 |
| 4. | − 1·2 | + ·1 | + 1·3 |
| 5. | − 1 | − ·1 | 0 |
| 6. | + 3·4 | + 4·4 | + 1·0 |
| 7. | + 3·7 | + 5·5 | + 1·8 |
| 8. | + ·8 | + 1·6 | + ·8 |
| 9. | 0 | + 4·6 | + 4·6 |
| 10. | + 2·0 | + 3·4 | + 1·4 |
| | Mean + ·75 | Mean + 2·33 | Mean + 1·58 |
| | S. D. 1·70 | S. D. 1·90 | S. D. 1·17 |

First let us see what is the probability that 1 will on the average give increase of sleep; i.e. what is the chance that the mean of the population of which these experiments are a sample is positive. $\frac{+\cdot 75}{1\cdot 70} = \cdot 44$ and looking out $z = \cdot 44$ in the

table for ten experiment we find by interpolating between ·8697 and ·9161 that ·44 corresponds to ·8873, or the odds are ·887 to ·113 that the mean is positive.

That is about 8 to 1 and would correspond in the normal curve to about 1·8 times the probable error. It is then very likely that 1 gives an increase of sleep, but would occasion no surprise if the results were reversed by further experiments.

If now we consider the chance that 2 is actually a soporific we have the mean increase of sleep $= \frac{2{\cdot}33}{1{\cdot}90}$ or 1·23 times the S.D. From the table the probability corresponding to this is ·9974, i.e. the odds are nearly 400 to 1 that such is the case. This corresponds to about 4·15 times the probable error in the normal curve. But I take it the real point of the authors was that 2 is better than 1. This we must test by making a new series, subtracting 1 from 2. The mean value of this series is + 1·58 while the S.D. is 1·17, the mean value being + 1·35 times the S.D. From the table the probability is ·9985 or the odds are about 666 to 1 that 2 is the better soporific. The low value of the S.D. is probably due to the different drugs reacting similarly on the same patient, so that there is correlation between the results.

Of course odds of this kind make it almost certain that 2 is the better soporific, and in practical life such a high probability is in most matters considered as a certainty.

*Illustration II.* Cases where the tables will be useful are not uncommon in agricultural work, and they would be more numerous if the advantages of being able to apply statistical reasoning were borne in mind when planning the experiments. I take the following instances from the accounts of the Woburn farming experiments published yearly by Dr Voelcker in the *Journal of the Agricultural Society.*

A short series of pot culture experiments were conducted in order to determine the causes which lead to the production of Hard (glutinous) wheat or Soft (starchy) wheat. In three successive years a bulk of seed corn of one variety was picked over by hand and two samples were selected, one consisting of "hard" grains and the other of "soft." Some of each of these were planted in both heavy and light soil and the resulting crops were weighed and examined for hard and soft corn.

The conclusion drawn was that the effect of selecting the seed was negligible compared with the influence of the soil.

This conclusion was thoroughly justified, the heavy soil producing in each case nearly 100 per cent. of hard corn, but still the effect of selecting the seed could just be traced in each year.

But a curious point, to which Dr Voelcker draws attention in the 2nd year's report, is that the soft seeds produced the higher yield of both corn and straw. In

view of the well-known fact that the *varieties* which have a high yield tend to produce soft corn, it is interesting to see how much evidence the experiments afford as to the correlation between softness and fertility in the same *variety*.

Further, Mr Hooker* has shown that the yield of wheat in one year is largely determined by the weather during the preceding harvest. Dr Voelcker's results may afford a clue as to the way in which the seed is affected, and would almost justify the selection of particular soils for growing seed wheat†.

The figures are as follows, the yields being expressed in grammes per pot.

| Year | 1899 | | 1900 | | 1901 | | Average | Standard Deviation | $z$ |
|---|---|---|---|---|---|---|---|---|---|
| Soil | Light | Heavy | Light | Heavy | Light | Heavy | | | |
| Yield of corn from soft seed | 7·85 | 8·89 | 14·81 | 13·55 | 7·48 | 15·39 | 11·328 | | |
| " " hard " | 7·27 | 8·32 | 13·81 | 13·36 | 7·97 | 13·13 | 10·643 | | |
| Difference ... ... ... | +·58 | +·57 | +1·00 | +·19 | −·49 | +2·26 | +·685 | ·778 | ·88 |
| Yield of straw from soft seed | 12·81 | 12·87 | 22·22 | 20·21 | 13·97 | 22·57 | 17·442 | | |
| " " hard " | 10·71 | 12·48 | 21·64 | 20·26 | 11·71 | 18·96 | 15·927 | | |
| Difference ... ... ... | +2·10 | +·39 | +·78 | −·05 | +2·66 | +3·61 | +1·515 | 1·261 | 1·20 |

If we wish to find the odds that soft seed will give a better yield of corn on the average, we divide the average difference by the standard deviation, giving us

$$z = \cdot 88.$$

Looking this up in the table for $n = 6$ we find $p = \cdot 9465$ or the odds are $\cdot 9465 : 535$, about $18 : 1$.

Similarly for straw $z = 1 \cdot 20$, $p = \cdot 9782$, and the odds about $45 : 1$.

In order to see whether such odds are sufficient for a practical man to draw a definite conclusion, I take another set of experiments in which Dr Voelcker compares the effects of different artificial manures used with potatoes on the large scale.

The figures represent the difference between the crops grown with the use of sulphate of potash and kainit respectively in both 1904 and 1905.

| | cwt. | qr. | lb. | | ton | cwt. | qr. | lb. | |
|---|---|---|---|---|---|---|---|---|---|
| 1904 | + 10 | 3 | 20 | : | + 1 | 10 | 1 | 26 | (two experiments in each year). |
| 1905 | + 6 | 0 | 3 | : | + | 13 | 2 | 8 | |

* *Journal of Royal Statistical Society*, 1907.

† And perhaps a few experiments to see whether there is a correlation between yield and 'mellowness' in barley.

The average gain by the use of sulphate of potash was 15·25 cwt. and the S.D. 9 cwt., whence, if we want the odds that the conclusion given below is right, $z = 1{\cdot}7$ corresponding, when $n = 4$, to $p = {\cdot}9698$ or odds of 32 : 1; this is midway between the odds in the former example. Dr Voelcker says 'It may now fairly be concluded that for the potato crop on light land 1 cwt. per acre of sulphate of potash is a better dressing than kainit.'

As an example of how the tables should be used with caution, I take the following pot culture experiments to test whether it made any difference whether large or small seeds were sown.

*Illustration III.* In 1899 and in 1903 "head corn" and "tail corn" were taken from the same bulks of barley and sown in pots. The yields in grammes were as follows:

| | 1899 | 1903 |
|---|---|---|
| Large seed ..... | 13·9 | 7·3 |
| Small seed ..... | 14·4 | 8·7 |
| | +·5 | +·6 |

The average gain is thus ·55 and the S.D. ·05, giving $z = 11$. Now the table for $n = 2$ is not given, but if we look up the angle whose tangent is 11 in Chambers' tables,

$$p = \frac{\tan^{-1} 11}{180^\circ} + {\cdot}5 = \frac{84^\circ\ 47'}{180^\circ} + {\cdot}5 = {\cdot}971,$$

so that the odds are about 33 : 1 that small corn gives a better yield than large. These odds are those which would be laid, and laid rightly, by a man whose only knowledge of the matter was contained in the two experiments. Anyone conversant with pot culture would however know that the difference between the two results would generally be greater and would correspondingly moderate the certainty of his conclusion. In point of fact a large scale experiment confirmed the result, the small corn yielding about 15 per cent. more than the large.

I will conclude with an example which comes beyond the range of the tables, there being eleven experiments.

To test whether it is of advantage to kiln-dry barley seed before sowing, seven varieties of barley were sown (both kiln-dried and not kiln-dried) in 1899 and four in 1900; the results are given in the table.

It will be noticed that the kiln-dried seed gave on an average the larger yield of corn and straw, but that the quality was almost always inferior. At first sight this might be supposed to be due to superior germinating power in the kiln-dried seed, but my farming friends tell me that the effect of this would be that the kiln-dried seed would produce the better quality barley. Dr Voelcker draws the conclusion "In such seasons as 1899 and 1900 there is no particular advantage in kiln-drying before sowing." Our examination completely justifies this and adds

"and the quality of the resulting barley is inferior though the yield may be greater."

| | lbs. head corn per acre | | | Price of head corn in shillings per quarter | | | cwts. straw per acre | | | Value of crop per acre in shillings * | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N. K. D. | K. D. | Diff. | N. K. D. | K. D. | Diff. | N. K. D. | K. D. | Diff. | N. K. D. | K. D. | Diff. |
| 1899 | 1903 | 2009 | +106 | 26½ | 26½ | 0 | 19¼ | 25 | +5¾ | 140½ | 152 | +11½ |
| | 1935 | 1915 | − 20 | 28 | 26½ | −1½ | 22¾ | 24 | +1¼ | 152½ | 145 | −7½ |
| | 1910 | 2011 | +101 | 29½ | 28½ | −1 | 23 | 24 | +1 | 158½ | 161 | +2½ |
| | 2496 | 2463 | − 33 | 30 | 29 | −1 | 23 | 28 | +5 | 204½ | 199½ | −5 |
| | 2108 | 2180 | + 72 | 27½ | 27 | −½ | 22½ | 22½ | 0 | 162 | 164 | +2 |
| | 1961 | 1925 | − 36 | 26 | 26 | 0 | 19¾ | 19½ | −¼ | 142 | 139½ | −2½ |
| | 2060 | 2122 | + 62 | 29 | 26 | −3 | 24½ | 22¼ | −2¼ | 168 | 155 | −13 |
| 1900 | 1444 | 1482 | + 38 | 29½ | 28½ | −1 | 15½ | 16 | +½ | 118 | 117½ | −½ |
| | 1612 | 1542 | − 70 | 28½ | 28 | −½ | 18 | 17¼ | −¾ | 128½ | 121 | −7½ |
| | 1316 | 1443 | +127 | 30 | 29 | −1 | 14¼ | 15¾ | +1½ | 109½ | 116½ | +7 |
| | 1511 | 1535 | + 24 | 28½ | 28 | −½ | 17 | 17¼ | +¼ | 120 | 120½ | +½ |
| Average | 1841·5 | 1875·2 | +33·7 | 28·45 | 27·55 | −·91 | 19·95 | 21·05 | +1·10 | 145·82 | 144·68 | +1·14 |
| Standard Deviation | — | — | 63·1 | — | — | ·79 | — | — | 2·25 | — | — | 6·67 |
| Standard Deviation ÷√8 | — | — | 22·3 | — | — | ·28 | — | — | ·80 | — | — | 2·40 |

* Straw being valued at 15*s.* per ton.

In this case I propose to use the approximation given by the normal curve with standard deviation $\frac{s}{\sqrt{(n-3)}}$ and therefore use Sheppard's tables, looking up the difference divided by $\frac{s}{\sqrt{8}}$. The probability in the case of yield of corn per acre is given by looking up $\frac{33{\cdot}7}{22{\cdot}3} = 1{\cdot}51$ in Sheppard's tables. This gives $p = {\cdot}934$, or the odds are about 14 : 1 that kiln-dried corn gives the higher yield.

Similarly $\frac{{\cdot}91}{{\cdot}28} = 3{\cdot}25$, corresponding to $p = {\cdot}9994$,* so that the odds are very great that kiln-dried seed gives barley of a worse quality than seed which has not been kiln-dried.

Similarly it is about 11 to 1 that kiln-dried seed gives more straw and about 2 : 1 that the total value of the crop is less with kiln-dried seed.

* As pointed out in Section V. the normal curve gives too large a value for $p$ when the probability is large. I find the true value in this case to be $p = {\cdot}9976$. It matters little however to a conclusion of this kind whether the odds in its favour are 1,660 : 1 or merely 416 : 1.

## SECTION X.

*Conclusions.*

I. A curve has been found representing the frequency distribution of standard deviations of samples drawn from a normal population.

II. A curve has been found representing the frequency distribution of values of the means of such samples, when these values are measured from the mean of the population in terms of the standard deviation of the sample.

III. It has been shown that this curve represents the facts fairly well even when the distribution of the population is not strictly normal.

IV. Tables are given by which it can be judged whether a series of experiments, however short, have given a result which conforms to any required standard of accuracy or whether it is necessary to continue the investigation.

Finally I should like to express my thanks to Professor Karl Pearson, without whose constant advice and criticism this paper could not have been written.

# SPLIT-HAND AND SPLIT-FOOT DEFORMITIES, THEIR TYPES, ORIGIN, AND TRANSMISSION*.

By THOMAS LEWIS, D.Sc., M.D. etc., and DENNIS EMBLETON, B.A., M.R.C.S. etc.

(*From University College Hospital.*)

CONTENTS.

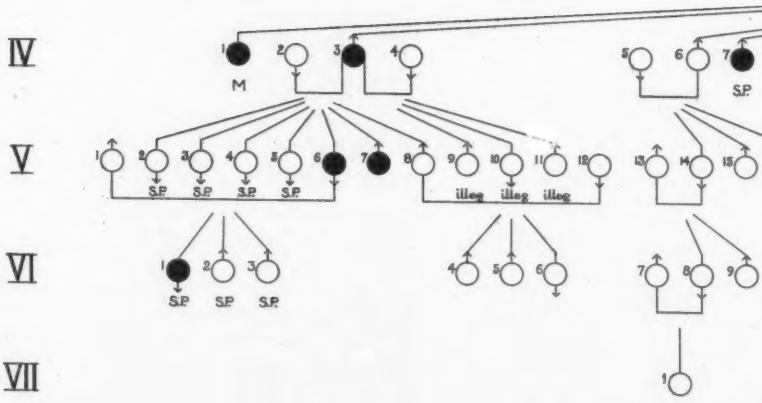
| | PAGE |
|---|---|
| Introductory | 26 |
| Detailed Account of the Deformities of the "G" Family | 27 |
| Types of Split-foot, their Terminology, and the Nature of the Cross-bones | 36 |
| Origin and Transmission of the Deformity | 43 |
| (*a*) General | 43 |
| (*b*) Maternal impressions | 44 |
| (*c*) Origin in an acquired lesion (extrauterine) | 44 |
| (*d*) Arrests of development | 45 |
| (*e*) Atavism | 45 |
| (*f*) Origin as a result of intrauterine conditions | 46 |
| (*g*) Origin as a "sport" | 48 |
| (*h*) Transmission of "Hereditary Split-foot," its stability, and the relation of its mode of transmission to Mendelism | 50 |
| Summary and Chief Conclusions | 56 |
| Bibliography | 56 |
| Explanation of Plates | 58 |

## *Introductory.*

THE deformity presented by the family described in this communication is one of exceptional interest and has so far received little attention. Accounts of it, of a brief and scattered nature, sometimes with figures, are to be found in the works of many of the well-known teratologists, such as Geoffroy Saint-Hilaire, Ammon, Förster and Otto.

In a Japanese temple an image of it has assumed, states Perthes (p. 136), the dignity of a god; a cloven hoofed deity.

* To this paper, written in November 1907, a short postscript, dated March 2nd, 1908, has been added. It will be found in the Miscellanea of this number.




No Consanguinity

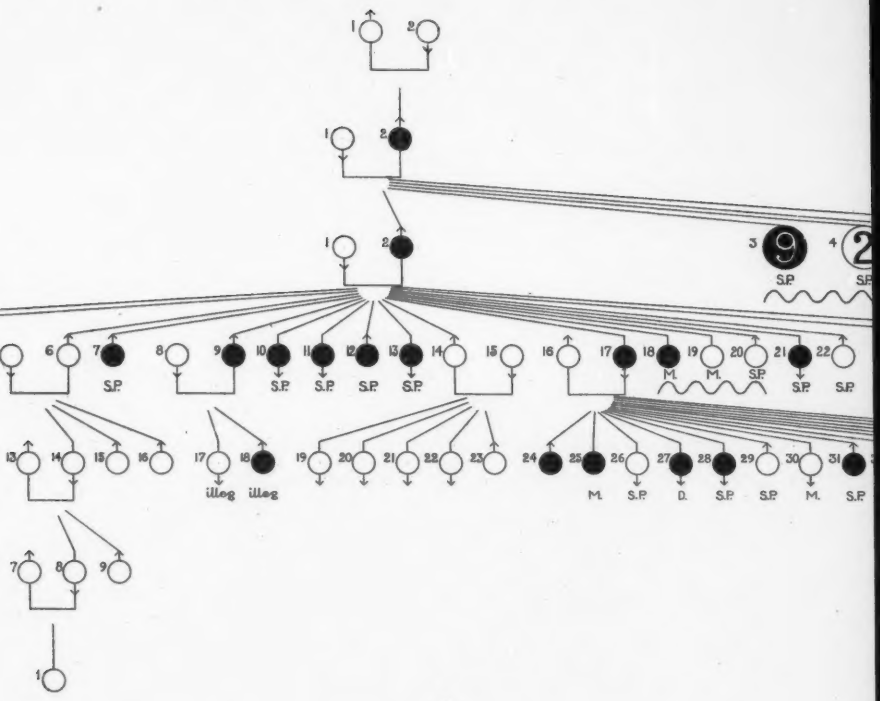
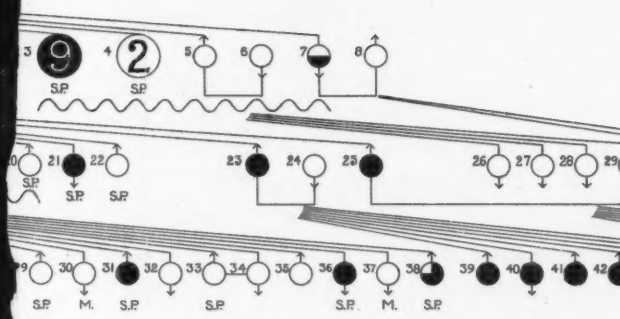
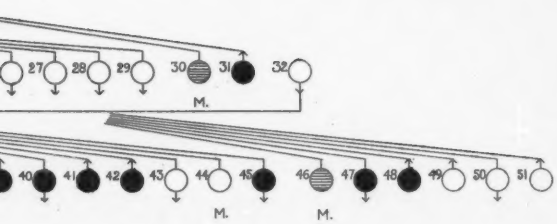



FIG. (I). GENEALOGICAL TREE OF THE "G" FAMIL


3
9
S.P.
4
2
S.P.
5
6
7
8
21
22
S.P.
S.P.
23
24
25
26
27
28
S.P.
30
M.
31
32
33
34
35
36
37
38
39
40
41
S.P.
S.P.
S.P.
S.P.
M.
S.P.

"G" FAMILY.

PLATE I.



The first collected description of cases which we have found is that by Kümmel. This author discusses the nature of "Spalt-Hand und Spalt-Fuss" at some length, but his account loses much of its value in that all types are indiscriminately mixed, and the question of heredity receives no attention. In a later monograph Perthes has brought a number of cases together on the same basis.

From the many varieties of split-hand and split-foot, one stands out prominently, and of this our "G" family presents notable examples. This type is characterised by its marked tendency to transmission, and by other features.

Including our family we have collected in all over 180 individual cases of the deformity, and have studied the details or detailed records of a large number of these*. We are therefore in a position to give a general sketch of the condition and its relation to general problems of heredity.

In searching the records for previously reported examples we have been fortunate in finding two earlier accounts of this remarkable family†, which contains more deformed members than any yet recorded. In 1886, before the introduction of "X rays," Anderson gave a description of it. A further brief summary was written by Tubby thirteen years ago. The addition of new members to the genealogical tree entitles the family to further consideration.

We must express our gratitude to Dr Higham Cooper of University College Hospital and to Dr Edward Shenton of Guy's for the invaluable aid they have given us in the screening and skiagraphy of the individual subjects of this investigation.

### *Detailed Account of the Deformities of the "G" Family.*

Observations have been for the most part confined to the bones of the hands and feet. Other deformities have been inquired into, and no such defects are known to exist. A complete examination of the family from this point of view has not been undertaken.

A thorough investigation of the soft parts of the hands and feet has also been neglected on account of the number of the deformed. From the observations made the following brief statements of the chief features may be made. The skin follows the lines of the underlying bones and tissues (Figs. 4, 5), and in no case has evidence of scarring been found, such as might suggest intrauterine injuries. The nails correspond in number almost without exception to the number of the terminal phalanges. When two fingers are closely syndactylised the nails are also united.

* Since this paper has been in the press, several other families showing the same deformity and living in London have been heard of, and other similar families in different parts of the country. An example which is typical is figured in Cruveilhier's well known "Atlas d'Anatomie pathologique, etc." The deformity appears to be a by no means uncommon one.

† Space does not permit of our giving the proofs that the three families are identical, but we are depositing private evidence with additional notes and radiograms, which future observers will find of value, in the library of the College of Surgeons.

The hands are curiously and variously misshapen and over the abnormal bony points thick masses of subcutaneous tissue are placed, which are frequently conica in appearance and as a rule situated towards the palmar aspect. These prominences are utilised in conjunction with the remaining finger or fingers for gripping. When a cross bone exists, the hand is far broader than might be expected from the number of metacarpals present. In no case is there any abnormality in the arrangement of the arteries as far as the wrist, and the tendons present correspond, so far as can be observed, to the presence of the metacarpals and phalanges to which they are normally attached. There are often contractures of the remaining fingers, so that they are flexed and drawn towards the middle of the hand. The contractures are dependent on the constitution and arrangement of the deeper structures and not upon the condition of the skin.

Each foot consists of two toes separated by a deep cleft. In exceptional cases there may be more than one complete toe in that portion of the foot bounding the outer side of this cleft, as evidenced by a double toe nail and longitudinal furrowing of the skin, but they are in no case separate. The cleft is always deep, and may have a length of four inches. As a rule it is two and a half to three and a half inches in the adult foot. The toes are widely apart at their centres, but are generally bent claw-wise at their extremities*. This separation of the toes gives the foot an unusual breadth in spite of the absence of metatarsals, a breadth which may amount to eight inches, or more. As a rule the inner border of the cleft runs parallel to the inner border of the foot for the greater part of its course and ends in furrows on the dorsum and sole of the foot. The outer border of the cleft is moulded over the remains of the metatarsal bones when these are present and is in consequence more irregular. On account of the inbending of the toe extremities to the middle line the foot is usually shortened. The tendons of the foot are arranged, like those of the hand, in conformity with the underlying bones; they are complete when the toes are present, and run to the ends of shortened digits.

The functional capacity of both hands and feet is great. The affected persons perform work which requires skilful manipulation. The needlework of some of the females is good; the handwriting of many is excellent. One individual is a boot-maker, two others drive cabs, none find difficulty in clothing themselves and most do so with extraordinary rapidity, even to the lacing of boots. The gait is normal; the toes are often opposible, more especially in the young, and frequently with considerable power and also delicacy. It is said of one that he is able in this way to lift pins from the floor with his feet. Nothing is more astonishing than the accurate and coordinate manner in which most of the functions of the normal hands and feet are carried out by parts so grossly deformed†. As to

* There is no evidence that this is produced by boot pressure, for it occurs in the younger subjects and has been observed in the new born. There is every possibility of its having been acquired *in utero*.

† The majority of those reporting this deformity express their surprise at the mobility of the malformed extremities.

the muscular structures which subserve these functions we can say little, as no opportunity of dissection has offered itself, but details may be gleaned from the few dissections figured by Otto, and from those reported by Schäfer.

In the following account of the defective conditions of the bones of the hands or feet the successive generations will be taken and the members of each generation described as they occur from left to right in the genealogical tree which has been given (Plate I, Fig. 1)*. In each case the source of information is stated, with the initials and where possible the dates of the birth and death of the individual. For the sake of brevity only those bones are described *which are present*. It must be understood that this applies rigidly unless it is otherwise expressly stated. The carpus which is omitted has in no case shown defect. The astragalus, os calcis and navicular bone have likewise never given evidence of affection. With regard to the last row of tarsal bones it must be said that we are often unable to render an exact account of it, and this for two reasons. First the frequent difficulty of identifying the bones by screening, and secondly the fact that when the cleft is deep considerable derangement of the tarsus occurs with separation of the bones and welding of them. The account is consequently confined so far as these are concerned to the chief points of interest observed. The omission of mention of carpal and tarsal bones must consequently not be interpreted in the light of our previous remark.

Generation I (I, 1 and 2)†. The parents of the first deformed individual are definitely stated by Anderson to have been normal. Enquiries from members of the family now living fully confirm this statement. They are said to have lived in Scotland‡.

Generation II, 2. Anderson states that J. G. was an only child, born in 1793 and dying in 1871. The descendants confirm this. The Brentford records show the dates to be correct. Several of the family now living knew him and give details of his deformity. The accounts correspond fairly well. The deformity of the feet was of the usual type but the feet were exceptionally broad. The hands, both grossly deformed, are stated to have had many fingers, including the thumbs.

Generation III. (III, 2.) J. H. G. (1825–1890). His children are unanimous in stating that he had one finger on each hand and two toes on each foot.

(III, 3 and 4.) For these two groups we are mainly indebted to Anderson, who states that all the children died in childhood. The family know very little of them, and the accounts from different members fail to correspond. Unfortunately the Brentford records go back to 1837 only, and the birth certificate of one W. G. born in 1838 is alone obtainable. We have therefore to rely on Anderson, and in consideration of the accuracy of his account in most points of importance in which we have been able to check it, we have little hesitation in accepting these figures. No information of the individual deformities is obtainable.

(III, 7.) M. P. (1832, still alive, married). This individual, who is 75 years of age, is feeble minded; visits to her abode have been without success, as both she and her friends steadily

* For description, see Explanation of Plates, p. 58.

† The bracketed figures refer throughout to the generation and number in the tree, Fig. 1.

‡ It must be stated that we have recently heard from Professor Pearson of the existence some seventy years ago of a Scotch family presenting "Lobster claw" deformity, said to date back to the seventeenth century.

refuse an interview. Her son's account agrees with that of Anderson. There is deformity of the feet only; each of these bears two digits.

Generation IV. (IV, 1, 18 and 19.) M. A. (IV, 17), who has lived with her mother for the last 17 years, states emphatically that there were three miscarriages in her generation. The first of these was deformed (IV, 1); and of the others, whose position in the family is uncertain, one was deformed the other undeformed (IV, 18 and 19).

(IV, 3.) H. C. G. (1845, still living, widower). Right handed. All four extremities screened. Anderson makes no mention of this man (cp. IV, 9 who also had illegitimate children).

*Right hand.* This is perfect with the exception that there is partial syndactyly of the 4th and 5th digits, and the terminal phalanx of the thumb is bent inwards.

*Left hand.* The last four fingers are normal. The 1st metacarpal is broad and its distal end is double. The inner facet bears a digit composed of three phalanges. The outer bears a small nodule of bone, the remains of an amputated digit which the man states was as long as its neighbour.

*Right foot.* The 4th and 5th toes have a normal complement of bones and are syndactylised throughout. The first toe shows no loss of bone. A deep cleft occupies the position of the absent 2nd toe. Between the cleft and the 4th toe lies the metatarsal of the 3rd toe, which has no phalanges. The cleft runs into the tarsus and the middle cuneiform appears to be absent.

*Left foot.* This is precisely similar to the right, but that it misses the terminal phalanx of the 4th toe.

(IV, 7.) Of this child we have no further information than that it was deformed and died in childhood; it is referred to by Anderson.

(IV, 9.) W. G. (approximately, 1852–1883). The hands and feet are stated by M. A. (IV, 17) to have been similar to those of her son W. H. A. (V, 24). Anderson makes no mention of this child, but the family have always been very reticent about him, as he had illegitimate children.

(IV, 10–13.) A. G. (approx. 1854–?), Ph. G. (1855–?). J. P. G. (1857–?), A. M. G. (1858–?). These four children are with one exception (IV, 10), mentioned by Anderson, and are said by him to have been deformed and to have died in childhood. Our enquiries make it probable that there were four deformed children lost at an early age, and the place of A. G. is taken in Anderson's account by a deformed son who died. No reliable details of the deformities are obtainable, apart from the fact that as usual all four extremities were affected.

(IV, 17.) M. A. (1861, still bearing children). Left handed. Skiagrams of all four extremities.

*Right hand*, Fig. 11. The thumb and 2nd finger are represented by the base of the 2nd metacarpal only. The 3rd metacarpal is normal, the 4th and 5th thickened*. From the apex of the 4th to the metacarpo-phalangeal articulation of the 5th a cross-bone extends, making a triple articulation from which a complete set of thickened phalanges arise.

*Left hand*, Fig. 10. The base of the 1st metacarpal is represented by a small nodule of bone. The 2nd metacarpal is thin at its distal end, the other three metacarpals are present. Between the extremities of the 3rd and 4th is a small bone articulating with the head of the 3rd and the cross-bone; it is the only representative of the phalanges of the outer three fingers. The cross-bone extends from the apex of the 4th metacarpal to the short 1st phalanx of the 5th. From the last articulation spring the terminal phalanges of two fingers, united throughout by bone. The corresponding nail is double.

* In the hands where the affection is mainly preaxial the postaxial metacarpals are always thickened, and in the following mention of it will be omitted.

*Right foot.* There are the usual inner and outer toes separated by a deep cleft. The outer metatarsal misses one phalanx, and those of the inner are united. At the bottom of the cleft is one piece of bone presumably the welded bases of the 2nd and 3rd metatarsals. The broad base of the 5th metatarsal indicates a fusion of it with the base of the 4th.

*Left foot*, Fig. 16. There are two toes only. The outer has an enormously thickened metatarsal articulating with the whole anterior surface of the cuboid. It bears three phalanges of which the last two are united. The inner metatarsal is also thick and bears two united phalanges. Its base articulates with the inner and middle cuneiforms. To the outer side of the cleft is a small piece of bone presumably the base of the 3rd metatarsal.

(IV, 20–22.) G. G., A. H. G., and Al. H. G. (born 1862, 1863 and 1865), are mentioned by Anderson, and our account confirms his; the three children died in childhood. No details of the deformity of A. H. G. are obtainable.

(IV, 23.) E. G. (approx. 1867, living, widower). Right handed. Skiagrams of all extremities.

*Right hand*, Fig. 23. All the metacarpals are present, the distal end of the 2nd is thin. The 1st, 4th and 5th have complete phalanges. The 3rd has a short 1st phalanx. The phalanges of the outer three fingers are united by skin.

*Left hand*, Fig. 22. This is similar except that the 2nd metacarpal is more fully developed.

*Right foot.* It is not possible to identify all the bones. On the inner side of the cleft is a large metatarsal (the 1st) bearing two phalanges. On the outer side two very heavy metatarsals occupying the articulatory positions of the 3rd, 4th and 5th. From the inner of these two, which articulate at their ends, springs a set of three phalanges.

*Left foot.* This is precisely similar.

(IV, 25.) I. G. (approx. 1869, living, married). Right handed. Screening of all four extremities.

*Right hand.* A small portion of the base of the 1st metacarpal is present. The other four are perfect. The 3rd bears the base of a 1st phalanx. The 4th bears the greater part of a 1st phalanx. The 5th has two complete phalanges, which are united to that of the 4th by skin.

*Left hand.* The thumb is not represented. The other metacarpals are perfect. The 3rd has the proximal half of the 1st phalanx. On the 4th is one and on the 5th two complete phalanges.

*Right foot.* Only two toes are present. The inner is composed of the complete 1st digit with the remains of the metatarsal of the 2nd united to its base. The outer toe consists of the complete 5th digit and the 4th metatarsal and its single 1st phalanx. A deep cleft runs as usual between the two toes and disturbs the arrangement of the tarsus. There is a piece of bone united to the base of the 4th metatarsal which appears to be the base of the 3rd metatarsal.

*Left foot.* There are two toes only. The inner composed of the complete 1st digit having its phalanges united; the outer consisting of a complete 5th digit, the bones of which are much thickened. Between the 5th metatarsal and the usual cleft lies a compound bone composed apparently of the 4th metatarsal united with its 1st phalanx and the base of the 3rd metatarsal. There is derangement of the tarsus but the base of the 2nd metatarsal appears to be represented.

(IV, 30.) The son (IV, 31) of M. P. (III, 7) states that there was a miscarriage before he was born, but no information as to whether it had any deformity is obtainable.

(IV, 31.) E. P. (approx. 1872, living and about to marry). Right handed. All four extremities screened.

*Right hand.* The 1st and 2nd digits are entirely absent. The 3rd metacarpal bears a small nodule of bone. The 4th and 5th digits are complete and syndactylised by skin.

*Left hand.* This is similar, it has in addition a thin 2nd metacarpal.

*The feet.* These are alike. The usual deep cleft occurs at the 2nd toe. The 1st and 5th toes are alone present and are complete.

Generation V. (V, 6.) A. S. (1870, living, married). All extremities screened.

*Right hand.* The last three fingers are normal. The 2nd misses the terminal phalanx, and the proximal phalanges are thin. The 1st metacarpal is present and bears two short deformed phalanges. From its inner side and about $\frac{3}{4}$ of an inch from its end, a small movable piece of bone springs. This is the remains of an amputated digit. There appears to be slight movement in the shaft of the metacarpal at this point. The amputated digit originally traversed the palm of the hand.

*Left hand.* The metacarpal bone of the thumb is broad at its distal end and bears on its inner facet three phalanges. The outer facet is unoccupied, but a scar indicates the position from which an additional digit was removed. It is said to have consisted of two bones and to have been as long as its neighbour. The rest of the hand is normal.

*Right foot.* The cleft occupies the position of the 2nd toe which is absent with the middle cuneiform. On the inner side of it lies the 1st metatarsal with two united phalanges. On the outer side lie a complete 5th digit and the metatarsal of the 4th, the phalanges of the latter and the bones of the 3rd toe are absent.

*Left foot.* This is the same but for the absence of union of the phalanges of the 1st toe.

(V, 7.) H. G. (approx. 1872, living, unmarried). Right handed. All extremities screened.

*Right hand.* The 1st metacarpal is small and weak; it has an extra phalanx. The 2nd, 4th and 5th metacarpals have complete sets of phalanges. The 3rd metacarpal has one complete and one incomplete phalanx. There is no syndactyly.

*Left hand.* The thumb is composed of metacarpal only. The bones of fingers 2, 3 and 5 are perfect throughout. From the 4th metacarpal springs a phalanx which is bifid distally, each facet bearing a complete set of phalanges. These sets are united by skin to the neighbouring fingers 3 and 5, but are themselves distinct up to the level of the common 1st phalanx. There are thus three fingers. That containing phalanges from metacarpals 4 and 5 has a single, and that containing phalanges from metacarpals 3 and 4 a double nail.

*Right foot.* There are the 1st and 5th toes, thickened and separated by a deep cleft. To the outer side of the cleft is the remains of a 4th metatarsal.

*Left foot.* This is similar.

(V, 18.) The illegitimate son of W. G. (IV, 9). Date of birth not known. He is said to have died. His hands are reported to have been similar to those of I. G. (IV, 25). The feet having each two toes.

(V, 24.) W. H. A. (1882, living, single). Right handed. All extremities screened, and left hand skiagraphed.

*Right hand.* This is very deformed and the identification of the bones is not quite certain. The arrangement is probably as follows:—The thumb is represented by a small nodule only. The remaining metacarpals are all present but the outer two are closely united by bone throughout There is a wide gap between the 3rd and 4th. The united 4th and 5th bear a common double 1st phalanx, which in turn bears two short single phalanges, articulated end to end. From the metacarpo-phalangeal joint of the combined 4th and 5th fingers a cross-bar runs to the head of the 3rd metacarpal with which it articulates.

Fig. 4

Fig. 5

FIG. 6 FIG. 7

FIG. 8 FIG. 9

FIG. 10

FIG. 11

FIG. 12

FIG. 13

Fig. 14

Fig. 15

Fig. 16

Fig. 17

Fig. 18

Fig. 19

Fig. 20

Fig. 21

FIG. 22 FIG. 23

FIG. 24 FIG. 25

*Left hand.* This is the most malformed of the series (Fig. 17). There are three heavy metacarpals articulating with the trapezium, os magnum and unciform bones. They probably represent the 3rd, 4th and 5th metacarpals. From the 4th springs a bone resembling a phalanx which articulates with the outer side of a welded mass of bone which itself springs from the head of the 5th metacarpal. The welded mass is 2 inches long and about 1 inch broad. It lies in the line of the metacarpal, from which it springs. At the free end of the mass are two terminal projections resembling short horns. The inner bears two small nodules of bone and a double nail. The outer has no nail. Between it and the inner are two small nodules of bone. There is thus evidence that the mass as a whole represents the portions of at least the inner three fingers.

*The feet.* There is one digit on each foot in the position of the 5th toe. Each bears three phalanges. The metatarsal at its base is thick and articulates with the whole of the anterior surface of the cuboid. The base of the 1st metatarsal is represented on both sides, and on the right probably the base of the 3rd also.

(V, 25.) Miscarriage, 4 months pregnancy. The mother states that the child had one finger and one toe on each limb.

(V, 27.) L. A. (1885, single, suffering from tuberculosis of the lungs). Right handed. Screening of all extremities.

*Right hand.* This is almost perfect but shows contractures, the thumb is small, the index finger is flexed. There is skin union of the 3rd and 4th fingers and slight bony union of the terminal phalanges. The bones are otherwise normal.

*Left hand.* The 1st metacarpal is surmounted by a small nodule of bone only. The 2nd metacarpal has one complete and a second incomplete phalanx. The remaining three fingers are normal. Between the 2nd and 3rd fingers is a small piece of bone about $1\frac{1}{2}$ cm. long, thin and narrow. It is on a level with the 1st phalanges and has the shape of a diminutive phalanx.

*Right foot.* On the inner side of the cleft the 1st toe is complete. On the outer side the inner toe is complete and to its inner side is a mass of bone obviously representing the 3rd and 4th metatarsals.

*Left foot.* On the inner side of the cleft the 1st toe is complete. On the outer side the 5th toe has metatarsal only. The 4th metatarsal has one complete and one incomplete phalanx, it is united at its base with the remains of the 3rd metatarsal.

(V, 28.) H. A. (1887– approx. 1895). Tubby states that this child had two fingers on each hand and two toes on each foot.

(V, 31.) J. A. (approx. 1892–1897). We are indebted to St Thomas's Hospital for notes of this child*.

*Right hand.* The thumb is not represented. The 5th finger has a full complement of phalanges. The 2nd, 3rd and 4th metacarpals are poorly developed and have no phalanges.

*Left hand.* Much the same as the right, but has two sets of phalanges for 4th and 5th fingers.

*Right foot.* Great and little toes only, of which the metatarsals are alone perfect. The cleft extends to the articulation between the metatarsal and the internal cuneiform.

*Left foot.* The same as the right, but the great toe has the stump of a metatarsal only.

(V, 36.) M. A. (approx. 1902–1905). The mother states that there was the usual deformity of the feet, each having two toes. The hands are said to have had one finger each.

* The notes are ward notes, there was no screening.

(V, 38.) J. A. (1906—1906). Screening of extremities. The child was born by the vertex, forceps were used and labour was prolonged. The after-birth was normal and there were no signs of adhesions or amputated digits. The post-mortem some months later was not obtainable.

*Right hand.* The five metacarpals were present though the 1st was small and bore no phalanges. The 2nd, 4th and 5th fingers had complete bones. The 3rd metacarpal had a phalanx double at its distal end and this in turn bore two complete sets of phalanges. The 2nd, 3rd and 4th fingers, namely four sets of phalanges, were united by skin (cp. Left hand of H. G. (V, 7)).

*Left hand.* This was normal.

*The feet.* The deformity was typical, the feet having each two toes, separated by wide clefts. The increased thickness of the 1st and 5th metatarsals and the general conformity of the foot to the type in other members of the family is of interest in that it indicates that little of the permanent adult deformity is due to use.

(V, 39.) E. W. G. (1886, living). Right handed. All extremities screened.

*Right hand.* Precisely similar to the left hand of R. E. G. which is figured (Fig. 7), except that its epiphyses are joined and that the 3rd metacarpal has no head.

*Left hand.* This is the same as the right but it has the head of the 3rd metacarpal.

*The feet.* The 1st and 5th toes are present with the full number of bones. The clefting is deep and the tarsus disturbed. There is welding of the remains of the 2nd and 4th metatarsals to the adjacent bones.

(V, 40.) E. M. G. (1888, living, single). Ambidextrous. Skiagraphy of all extremities.

*Right hand*, Fig. 13. There is no trace of the bones of the thumb. The remaining metacarpals are present. The 3rd bears a half phalanx. From the distal end of the 4th a large complex cross-bone extends to the head of the 5th, articulating with it and the conjoined bases of the two phalanges, presumably those of the 4th and 5th fingers. The united phalanges are separate at their distal ends, and bear, the inner one, the outer two phalanges, the latter united. We are inclined to the view that the 1st phalanx of the 5th finger is represented in the inner end of the cross-bone.

*Left hand*, Fig. 12. The base only of the 1st metacarpal is present. The 2nd metacarpal is thin and like the 1st bears no phalanges. The remaining metacarpals are present and of them the 4th and 5th have the full complement of phalanges, the two sets syndactylised by skin. From the head of the 3rd metacarpal a curved cross-bone proceeds to the metacarpo-phalangeal joint of the 4th finger.

*Right foot*, Fig. 15. On the inner side of the cleft is the 1st metatarsal bearing two phalanges. On the outer side are the 4th and 5th metatarsals, carrying between them two sets of united phalanges. The clefting is deep and causes some displacement and modification of the tarsus.

*Left foot.* This is practically identical; the middle cuneiform appears to be welded to the base of the 4th metatarsal.

(V, 41.) R. E. G. (1890, living). Right handed. Skiagraphy of all extremities.

*Right hand*, Figs. 4 and 7. The thumb is entirely absent. The 2nd metacarpal misses its distal end, the remainder are complete. The 5th finger is complete, the base of the 1st phalanx has a double articulation, the one facet for the metacarpal the other for the cross-bone which extends from this joint to the head of the 4th metacarpal.

*Left hand*, Figs. 4 and 6. This is similar, except that the 2nd metacarpal is represented by its base only.

*Right foot*, Figs. 5 and 9. There are the usual two toes and deep clefting. The outer toe has two phalanges, the inner two irregular ones. The base of the 4th metatarsal is represented.

The middle cuneiform appears to be absent, while the external seems to be welded to the cuboid.

*Left foot*, Figs. 5 and 8. This is similar, but the base of the 4th metatarsal is welded to that of the 5th.

(V, 42.) J. T. G. (1892, living). Right handed. Skiagraphy of all extremities.

*Right hand*, Fig. 19. The thumb is not represented. The four outer metacarpals are present, the 2nd being thin. The 4th and 5th have complete sets of phalanges. A cross-bone proceeds from the metacarpo-phalangeal joint of the 4th finger to the head of the 3rd metacarpal.

*Left hand*, Fig. 18. The five metacarpals are present. The 1st has one short phalanx. The 2nd none. The 3rd a half phalanx, replacing the cross-bone of the opposite side. The 4th and 5th have complete phalanges, syndactylised by skin.

*Right foot*, Fig. 25. All the metatarsal bones with the exception of the 2nd are present, the 2nd is replaced by the cleft. The 1st metatarsal bears two phalanges. The 4th and 5th have a common set.

*Left foot*, Fig. 24. To the inner side of the cleft lies a complete 1st toe. The 2nd is entirely replaced by the cleft. The base of the 3rd is welded to the 4th, and scarcely appears. The 4th is complete and has an irregular phalanx. The 5th has a complete set of phalanges.

(V, 45.) L. V. G. (1897, living). Right handed. Skiagrams of all extremities.

*Right hand*, Fig. 14. The thumb is not represented. The 2nd metacarpal has no distal epiphysis. The 4th and 5th have complete phalanges which show skin syndactyly. A cross-bone runs from the 4th metacarpo-phalangeal joint to the head of the 3rd complete metacarpal.

*Left hand.* This is similar but the 1st metacarpal has a short representative.

*Right foot*, Fig. 21. The cleft occupies the position of the 2nd metatarsal and the middle cuneiform, which are absent. The 1st toe is complete. The 3rd toe is not represented. The proximal half of the 4th metatarsal is present, and the 5th toe is complete.

*Left foot*, Fig. 20. The cleft occupies the same position. On its inner side is a complete 1st toe. On its outer side are the base of the 3rd metatarsal, the whole 4th metatarsal and the complete 5th toe.

(V, 46.) Miscarriage. The mother does not know if it was malformed.

(V, 47.) K. G. (approx. 1901, living). Left handed. All extremities screened.

*Right hand.* The 1st metacarpal has its base only, the others are normal. The 2nd and 3rd metacarpals bear each the base of a 1st phalanx. The 4th and 5th have complete phalanges which are syndactylised throughout by skin.

*Left hand.* The thumb is complete and has an extra phalanx. The 2nd finger has the metacarpal, 1st and part of the 2nd phalanx. The 3rd, 4th and 5th fingers have the normal complement of bones, those of the 4th and 5th being joined by skin.

*Right foot.* Toes 1 and 5 are present, their bones which are complete are much thickened. Between the cleft and the outer toe is a mass of bone representing probably the base of the 4th metatarsal. The tarsus appears to be unaffected.

*Left foot.* As usual there are two toes with a deep cleft between them. These are mainly composed of the thickened bones of the 1st and 5th digits. Between the cleft and the outer digit are the remaining three imperfectly developed metatarsals.

(V, 48.) W. G. (approx. 1903, living). Right handed. All extremities screened.

*Right hand.* The base of the 1st metacarpal is represented by a nodule only. The outer four metacarpals are normal. The 4th and 5th metacarpals bear complete sets of phalanges, syndactylised by skin.

*Left hand.* The hand is similar to the last, no part of the 1st metacarpal is seen however, and the 2nd metacarpal bears a short 1st phalanx.

*Right foot.* The foot is cleft at the 3rd metatarsal which is absent. The 1st and 5th toes are complete, and there are two masses of bone lying between these, one on each side of the cleft and representing apparently the remains of the 2nd and 4th metatarsals.

*Left foot.* This is practically identical with the right.

Generation VI. (VI, 1.) M. A. S. (approx. 1890, died a few weeks after birth). According to the mother, the right hand had an extra thumb, and the left had some slight abnormality of which no trustworthy account is given. The feet are said to have had two toes each with deep clefts between.

In the above description the normal individuals have not been included. A number of them have been seen and the reports of the family in respect of them have been confirmed. For th families of T. G. (III, 5), J. F. W. G. (IV, 6), and T. C. G. (IV, 14) we have had to rely on the statements of the remaining members, and have cross-examined several separately and so confirmed the statements with regard to them. The family of J. F. W. G. is no longer in England, and the present whereabouts of the other two families is not ascertainable.

## *Types of Split Hand and Foot, their Terminology and the Nature of Cross-bones.*

With those cases of split hand or foot which are split by reason of their duplicity we are not concerned. Such cases have eight or ten digits on the affected extremity and are rare (see Murray and Giraldes). Our main purpose is the differentiation of the scattered forms of ectrodactyly, which yield a cleft hand or foot, from the hereditary and symmetrical type.

The majority of these irregular types (examples of which are referred to in Section C of the Bibliography), have been reported as single cases, some without reference to heredity, others with a definite history of normal parents. Kümmel has reported a family in which five individuals were deformed, by a splitting of one or both hands. At the same time he analyses a number of cases of split-hand and split-foot but without separating the types. In the cases which have been enquired into, the split usually occurs at the 3rd finger, and depends on the absence of the corresponding bones. It may be more extensive and include the 2nd and 4th fingers. Only one case of single split-foot, unassociated with other lesion, and in any way comparable to those of the "G" family, has been found. It is that quoted by Kümmel after Gintrovicz, in which the phalanges of the 3rd toe were absent. Further cases are figured in the older works in which anomalous forms of split-foot are associated with phocomelia and other extensive deformities.

When these irregular types are eliminated there remains a by no means uncommon deformity of which the "G" family presents notable examples. There are over 180 cases* of it now on record, in all but 13 of which there is a family history of similar defects. It is characterised by symmetrical clefting of the feet,

* In the following pages our account depends on an examination of about 100 cases of the deformity, in reports or in life.

with complete syndactyly of the remaining two groups of toes, and an irregular though often symmetrical deformity of the hands. Its marked tendency to become hereditary will be subsequently discussed.

Dealing first with those cases which have been reported as individual members of deformed families (see Bibliography, Section A), and including our own with them, the following conclusions may be drawn. All four extremities show deformity in the vast majority of cases. In a few one hand and both feet are affected[9] (V, 38)*, in a slightly greater number of instances the feet are alone abnormal[9, 11] (III, 7). The hands are never defective in the absence of foot deformity†. One foot is never malformed alone. The feet tend to illustrate one of three main types of malformation. Of these the commonest is that in which all metatarsals except the 1st and 5th are absent (Figs. 8, 9), or rather the 1st and 5th metatarsals are alone definitely represented. In the second type the main deformity falls on the 2nd toe (Figs. 20, 21, 24, 25), in the third type on the 3rd toe (Figs. 15, 16). The last two types are of about equally common occurrence‡. Although these main types may be isolated intermediate conditions remain. For descriptive purposes it may be said that, starting with the 2nd or 3rd toe as a basis, the lesion is extended§ to the adjacent toes in the order of their proximity. First the phalanges and then their metatarsals disappear. When one toe is completely absent in all its parts, which is the rule, the neighbouring phalanges are rarely if ever unaffected. Lastly this exception must be added, that the 1st toe tends in the vast majority of cases to escape. In a few instances this toe also is shortened‖, and in some has been absent[3, 16, 17] (V, 24 and 31). In Perthes' case there was, in addition, affection of the corresponding cuneiform, navicular bone and astragalus, with some loss from the lower end of the tibia, but his case stands alone in this respect. Absence of the 5th toe must be equally rare. It may be shortened by the loss of one or more phalanges. Also it has happened that one or more phalanges have appeared to be absent when the 4th toe seemed to have its full complement (IV, 23), but it is doubtful if in these cases the phalanges have not been displaced. It is an almost rigid rule that a toe has fewer bony representatives according as it approaches the centre of the cleft. When the two outer toes remain intact they may be syndactylised by bone, but in any case they are joined by skin. There are frequently remains

* The small number refers to the bibliography, the bracketed figures to the individuals in Plate I, Fig. 1.

† This statement is made somewhat reservedly. There is no reported case in which the hands are described and the feet are stated to be normal. If it does occur it must be quite exceptional.

‡ In comparing the frequency of types, the tendency in some families for one type to predominate, and the number of deformed members in each family have been taken into account. The examples referred to (Figs. 15 and 16) are not so decisive as many which occur in other families.

§ In speaking of the extension of the deformity it must be clearly understood that the word is used purely for descriptive purposes and is in no way intended to imply that such extension has actually taken place.

‖ On the other hand it may be lengthened in exceptional cases by reason of the presence of three phalanges.

of tarsal bones and metatarsals in the cleft, but their welding together and disorder often preclude an accurate description, compatible with brevity (Figs. 15, 16, 21). It is an invariable rule that a proximal bone is never absent when a corresponding distal one is present. Of the metatarsals, which remain, those bearing terminal phalanges are usually considerably thickened while those which bear no phalanges show hypoplasia, which is greater peripherally. There are often indications that the thickening of the 1st and 5th metatarsals is due to welding with the remaining portions of the neighbouring metatarsals, namely the 2nd and 4th (cp. Figs. 8, 9, 15, 16).

From the detailed description of the "G" family and this general description it will be seen that there is a good deal of variation in the feet, but at the same time the points of resemblance are still more striking. In short these resemblances are the clefting with the absence of at least all the phalanges and half the metatarsal of one toe and the absence of one or more phalanges from neighbouring toes (Fig. 25), the syndactyly of all the remaining toes, or parts of them, into two groups, the broadening of the feet and the inturning of the terminal phalanges, and lastly the presence of split foot on both sides. *Briefly the variation in the deformity is of degree and not of kind.*

Passing to the hands, it will be remembered that they are not affected unless the feet are split. Also that in rare cases, one hand may be normal[9], or more often both normal[9, 11]. Further they may show syndactyly or polydactyly without other defect of the bones (IV, 3, etc.). There may be loss of one or more phalanges from one finger. Finally gross deformity may be found, and this is the rule. This again shows great variations but two general types may be separated. *In one* the hand is split, by reason of ectrodactyly of the 3rd (middle) finger. Such splitting commonly extends to, but not into the carpus*; on the other hand it may fail to involve the metacarpal (V, 7), and in this case the corresponding phalanges may constitute cross-bones. In a single instance[9] there has been a strong suggestion of division of the 3rd phalanges, the component parts lying to either side of the gap. As in the feet the process may be more extensive and involve the 2nd (Fig. 22) and 4th digits. *In the other* variety the ectrodactyly proceeds† from the radial side and the bones escape according as they approximate to the ulnar side. In this form it is usual for the 5th (Figs. 6, 7), or 4th and 5th (Figs. 14, 18, 19) fingers to have a complete complement of bones. Exceptionally no finger is complete. In two instances the 5th finger has shown more change than the 4th[17, 21]. The 3rd metacarpal is rarely and the 2nd usually affected; frequently they both carry one or more phalanges, those of the outer being more complete. The thumb is in most instances absent or represented by a short bone only, but it may show any grade from this up to the possession of all its bones (Fig. 23).

* Affection of the carpus has been thought probable by writers [28 etc.] in certain cases, but no clear evidence of it has as yet been forthcoming. Through the kindness of Professor Pearson we have recently seen a case, as yet unpublished, in which it appears to occur.

† Compare footnote on page 37.

There are thus connecting links between the two main types (cp. hand of V, 7). The type in which there is preaxial affection as opposed to splitting is a little commoner. In neither type is the radius or ulna ever affected, and it can only be in exceptional cases, if in any, that the carpus suffers. Both types may be associated with syndactyly or polydactyly.

Syndactyly is generally confined to the 4th and 5th digits (Fig. 14), but when more are present and complete it may unite these; the rule is for two adjacent fingers to be united by skin when they have phalanges. The union may be by the deeper structures as well, and in a fair number of cases is bony* [11, etc.]. The subject will be reverted to under the discussion of the cross-bones.

Polydactyly occurs less frequently. The supernumerary digits may be rudimentary[5], the extra digit may be articulated to the end of a phalanx (V, 38), to the end of a metacarpal (V, 6, 7), or to the carpus[30]. It may be fixed to the side of the thumb metacarpal (V, 6). When articulating with the end of a metacarpal or carpus the digit is fully developed and has three phalanges, even when the joint is between it and the 1st metacarpal. Thus the varieties of polydactyly which occur are similar to those found when unassociated with other lesions (cp. Annandale). It occurs most commonly in those split foot cases where there is no other affection of the hands but is not confined to these (V, 6). There is only in isolated cases more than one extra digit on each hand†.

Thus it is seen that the hands show even more variation than the feet; nevertheless they show points common to the majority. Notably the mobility, the fleshy pads over the abnormal bony prominences, contractures and syndactyly of remaining fingers, and the exemption of the carpus. When the two main types are taken separately the bony deformities in different individuals show very close resemblances, as detailed above; moreover there are connecting links between these main types.

It has now been shown how much both hands and feet vary, but that in spite of this *there are points of strong resemblance between the affected limbs of different individuals.* Moreover there is a variation, usually limited, on the two sides of the body and the variations of hand or foot appearing in different families are strikingly similar. Thus in the larger families "G," [5, 9, 11‡], though there may be a predominant type, *yet it is the rule for each family to provide examples of many types, and in one family almost all types are represented*[11]. Again it has not been found possible to connect the occurrence of any particular type of foot with any particular type of hand. It was at first thought that the type of foot in which the split occurs at the 2nd digit might be more often found in association with that type of hand in which the preaxial side is mainly defective, and that splitting of the feet and hands at the central digit might be found to occur more commonly in conjunction with each other. But when allowance is made for the tendency shown here and there

* In these the union is closest at the distal end, and the nails are united in a central furrow.

† For exceptions see [5, 18].

‡ "G" signifies the family now reported.

for the transmission of a particular type these relations do not hold. *With few exceptions each type of foot is found in association with each type of hand.* Lastly, although in certain families, as in that of Mayer, such a combination as split hand and split foot occur together and predominate, yet there are *families in which many combinations of types are blended.* Dealing with the main facts, it may be said that when this quadruple deformity, which presents extreme variation of form, exists as a hereditary malformation, the constantly occurring similarity between individual parts and individual combinations, when taken into account with the common tendency to transmit along definite lines, leads to one conclusion, *namely that in all these families the deformity of each individual has the same fundamental basis; in fact that the deformity is a distinct entity.*

It now remains to bring into line those cases which have been reported as isolated cases (Bibliography, Section B). At first sight the absence of the hereditary history sharply breaks them off from the familial cases so far discussed. There are in all, however, only 13 cases of which we can find descriptions or figures, and of these eight are not dealt with from the hereditary standpoint at all. In five it may be assumed that the parents were normal[23, 25, 26, 28, 31], but two of these were children and the third a lad. In some of the 13 cases there is consequently a possibility that the question of heredity was neglected, in others there was no opportunity of studying it. In yet other cases the absence of hereditary history may have been due to the reticence of the patient; from our experience of the "G" family we may say that had we first met certain members of the family now known to us, we should have had great difficulty in extracting the necessary information. Illegitimacy, several examples of which we have ourselves met, provides another possible source of error. Again a certain proportion of the 13 cases are likely to have been the first in their family to show deformity*. Lastly it is more than possible that a deformed person may have the potentiality of transmitting and still fail to do so. These considerations lead us to believe that some if not all of the isolated cases either actually belonged to deformed families or were by no means lacking in the power to transmit the malformation, and there is consequently no reason to exclude them upon this score. Moreover, considering their deformities, *no isolated case occurs in the list which is given which does not find its almost exact counterpart in members of the families given in section A.* Consequently as there is little hesitation in regarding the familial cases as examples of the same entity so it is considered that the isolated cases, though varying widely, are of the same nature and may be classified with them as illustrating a deformity *sui generis.*

Having now, as we believe, established† the deformity under discussion as an entity, the terminology may be briefly referred to. So many terms are found in

* The proportion of original deformed members to those occurring in the 19 families is approximately as 1 is to 8.

† It is not possible to bring forward the complete evidence in support of this view, as space does not permit, but those who are acquainted with the reports and figures cannot fail to be struck by the close similarity of separate cases. The position is in reality far stronger than it can be made in a general outline sketch.

the present nomenclature that confusion is unavoidable. "Perodactyly," "Peromanus," "Peropod," "Perochirus," "Ectrodactyly"; "club-foot," "split-foot," "crab or lobster claw," and the foreign equivalents "Hummer-Schere," and "main ou pied en pince de homard" are a few of its more simple designations. Many of these terms are applied to a variety of malformations and none are sufficiently comprehensive as descriptive terms. We are not concerned with any other class of split-hand or split-foot but that of which the "G" family presents examples, and in consideration of the constant and fundamental character of the foot lesion*, the term "Hereditary Split-foot" is proposed for this deformity†. The term is perhaps insufficient in that it contains no reference to the defect of the hand, but the variability of the latter is so great that no adequate term is possible without lengthening it to an undesirable extent; moreover there are cases in which there is no defect of the upper extremities.

*The nature of the cross-bones.* This is a subject of very considerable interest, and particularly in the "G" family. What may be termed the clean form is illustrated in Figs. 6, 7. These straight bars of bone‡ have shafts with concave borders and end in articular surfaces, each of which may in young subjects be shown to have developed from a special epiphyseal ossification centre. In older subjects there are indications of the same. A cross-bone has not been met with in a sufficiently young hand to enable us to identify the centres of the shaft, but in some there is an indication of its double origin (Fig. 12). The outer articulation is a triple one in which the base of the 1st phalanx and the head of the metacarpal bone take part. The cross-bone may bridge over the space between the heads of the 4th and 5th (Fig. 6), or the 3rd and 4th (Figs. 14, 19) metacarpals. The outer extremity of the cross-bone has two facets, the inner extremity one facet, which may or may not be in contact with the head of the metacarpal corresponding; in the latter case the articulation is with the side of the head or shaft of the cross-bone. The type illustrated in Figs. 6 and 7 occurs in the hand of IV, 17 (Fig. 11) and in both the hands of V, 39 and 41, the children of IV, 23. Articulating with the 3rd and 4th metacarpals it occurs in both hands of another member of the same family, namely V, 45 (Fig. 14), and in one hand of V, 42 (Fig. 19), and V, 24. These two types have been found in no other members of the "G" family. The irregular forms occur in IV, 17's right hand (Fig. 10) and in both hands of V, 40 (Figs. 12, 13). Thus cross-bones are limited in this family, so far as we are able to judge, to IV, 17, her eldest child, and her brother's children, but from the description of their father's hands (III, 2) it appears probable that he was similarly deformed, and also other children of IV, 17

* The constancy of the foot deformity as opposed to that of the hands may possibly be associated with the phylogenetic evolution of the extremities. The feet being considered as undergoing retrogressive and the hands progressive development.

† The term has the additional advantage of being in unison with the German appellation "Spalt-Fuss."

‡ Which are now recorded for the first time.

herself. Of the five deformed children of IV, 23, four have cross-bones in each and the fifth in one hand, and this in spite of the fact that the father's hands are free from it (Figs. 22, 23).

Cross-bones of a somewhat different character have been reported by many writers[8, 9, 11, 12, 13, 46 etc.], and have been discussed by Perthes. Such cross-bones are obviously displaced phalanges, lying across the head of the 3rd metacarpal or articulating with phalanges of the second row. The method in which such cross-bones, as are figured in v. Bergmann's *Surgery* (Figs. 183, 184), are formed is strictly comparable to the method of production of a distal bony syndactyly. The hand shown in Fig. 10 shows a very similar condition, though more advanced. In this figure though the 1st phalanges appear to be united, yet when screened there was a doubt about it. The main cross-bone in this figure can only be interpreted as the 1st phalanx of the 4th finger, and the shorter piece, perhaps united with it, as the corresponding phalanx of the 5th. A similar explanation will not apply to the type shown in Figs. 6 and 7, for complete sets of phalanges spring from the triple joints. Moreover the presence of two epiphyses precludes the possibility of interpreting the cross-bone as one phalanx. A further explanation must consequently be sought. Now there can be but two possibilities, for either the cross-bone is an entirely new formation or it is developed from the representatives of parts normally present. In regard to the last there are two explanations, the cross-bone may represent phalanges or metacarpals. The latter is excluded by such hands as the left of V, 45, in which all the metacarpals are present. The cross-bone may thus be said to represent either united phalanges or to be an entirely new formation. If the former view is adopted the phalanges concerned must necessarily be those normally situated further to the radial side of the hand than the position of the cross-bone itself; for instance in the right hand of V, 45 (Fig. 14) they should come from the 2nd and 3rd digits, and in V, 41 (Fig. 6) from the 3rd and 4th. There is no difficulty in the way of accepting this displacement, for such cross-bones are the only possible representatives of the phalanges corresponding to those fingers which lie between the space bridged by the cross-bone and the thumb (compare the hand of V, 40 (Fig. 13) and the subsequent explanation of her deformity). The view that the cross-bone is developed from ossifying centres which normally represent phalanges is by far the most plausible and is further supported by the following considerations. In all the cross-bones there is a strong resemblance between the extremities of the bones and the bases of 1st phalanges. In the right hand of V, 40 (Fig. 12), the cross-bone is bent and indicates such a union. In the hands of V, 42 (Figs. 18, 19) a cross-bone on the left side (Fig. 19) is replaced on the right (Fig. 18) by the proximal half of a 1st phalanx. Lastly there can be no doubt that in the cross-bones shown in Figs. 10 and 13 phalanges are represented, and when this is taken into consideration with the fact that cross-bones in other families can only receive rational explanation in this way further evidence is hardly wanting. Our conclusion is that in the hands of V, 41 (Figs. 6, 7)

the cross-bone represents the united bases of the 1st phalanges of the 3rd and 4th fingers, welded together and displaced inwards*.

We now come to the hand of V, 40 (Fig. 13). In this the presence of the base of the 1st phalanx of the 3rd finger precludes the possibility of displacement†; as a matter of fact such a supposition is unnecessary, for the cross-bone may be said to represent the 1st phalanges of the 4th and 5th fingers. The first being horizontally and the latter vertically placed and much thickened. The U-shaped bone is the representative of the 2nd row of phalanges. It is noteworthy that the terminal phalanx of the 4th finger shows an indication of being double, but however the deformity is explained the same difficulty will arise. The deformity of the hand is so gross that so small a factor is of little consequence.

In concluding this section it is worthy of note that although the cross-bone formation is so closely allied to the process of bony syndactyly, and that this occurs in hands and feet, yet cross-bones are never observed in the lower extremities.

### *Origin and Transmission of the Deformity.*

(*a*) *General.* In entering upon a discussion of the origin and hereditary transmission of a human deformity, there are several considerations of importance. But before discussing them it is desirable to lay particular stress upon one fact, namely that the origin of the deformity in the first member of the family affected may be—but is not necessarily—a question entirely apart from that of its transmitted origin in later members of that family. Confusion of the two problems arises so frequently in medical communications that this emphasis is imperative.

Comparing the standpoint of the clinician or pathologist with that of the vegetable teratologist, it will be seen that the former labours under distinct disadvantages. For the latter is aided by the prolific nature of the material with which he works, a vast number of progeny arising from a single parent and one generation following rapidly on another. Moreover the material is under control; not only is he able to restrict the complicating factors by inbreeding, but from the hermaphroditic nature of much of the material he has the opportunity of tracing hereditary characters in a stock derived from a single ancestor. Those who investigate the smaller, more fertile and quickly breeding species of the animal kingdom, enjoy similar but more restricted advantages. The human teratologist on the other hand must work entirely with the mixed material provided by the irregular intermingling of a community. Rarely has the investigator of a human family the opportunity of examining three generations, and often but two or one; his information has therefore to be drawn to a large extent from hearsay evidence, or from the old and oft-times misleading accounts of other members of the family.

* Corresponding explanations apply to the other bones of this type. A rational elucidation of the hand shown in Fig. 17 has not been obtained.

† There is no reason to believe that an actual displacement ever takes place, but that the bone is developed *in situ* is certain from the nature of the articulations.

It is of first importance, from the point of view which we are considering, that in reports of family affections the whole genealogical tree be investigated as accurately as circumstances permit. Many of those who have written on deformities of the hands and feet have contented themselves with an account of affected members only. The tendency of members of the later generations to be acquainted only with the deformed relatives preceding them is very great and the fallacies arising from it must be carefully guarded against*. It is possible that certain cases in which dominance of deformed members is in great excess in the earlier generations may be accounted for to some extent in this way.

Although clinical observation may fail to contribute so richly to the questions at issue in dealing with hereditary problems, as do observations on the lower grades of animal and plant life, and though to a large extent laws formulated from the latter must serve as a basis in the discussion of similar problems in man, yet the study of hereditary deformities forms a by no means unproductive field of research. And the harvest from it remains at present almost totally ungathered.

The principles governing transmission in man are assuming daily an aspect of greater importance, both from the academic and practical standpoint, and there is every prospect that their elucidation will be of vital importance in the near future.

(*b*) *Maternal impressions.* In the literature of deformities of the extremities a history of maternal fright occurs with great constancy. It appears to be as great a matter of indifference, as to in what period of gestation the "impression" was produced, as it is of little consequence whether the child's deformity is identical with the object which is attributed as its cause.

In connection with our family, Anderson originally stated that the condition was attributed to a packet of lobsters. The legend is now content to impute it to a crab. The whole subject would be unworthy of mention were it not for the fact that the superstition is not only firmly rooted in the minds of the lower classes, but appears to find favour with certain medical writers[5, 61]†. It is to be explained in great measure by their want of recognition of common teratological defects and monstrosities. We have no intention of dealing further with a theory which has been satisfactorily interred by numerous writers, but content ourselves by referring to the remarks of Förster, Vrolik and Lewis.

(*c*) *Origin in an acquired lesion* (*Extrauterine*). The old controversy of the Lamarkian doctrine, as opposed to the antagonistic theory of the exclusive transmission of "in-born" characteristics, is one which it would be out of place for us to enter upon at any length.

* To those who have never undertaken the task of compiling a genealogical tree of a family of working class people the difficulties may not be apparent. Taking the "G" family, it is the rule for the adults to be unaware of the names of many of their brothers and sisters and it sometimes happens that the names and number of the children are unknown to a parent.

† The possibility is discussed at some length by some writers, who have mentioned the possibility of transmission by recurring "impressions" from the children (compare [11]).

Both the family of Kümmel* and that of Scoutetten arose from a parent with an acquired defect of the limbs, and the last author actually attributes the defect to this source. He has been criticised for so doing by Perthes. It is almost needless to say that there was no evidence that the deformity of the offspring was in any way comparable to that of the parent. With this solitary example, which has been successfully controverted, there is no particle of evidence in a single case to show that such transmission has taken place. The supposition can be dismissed.

(*d*) *Arrests of development.* There would be no occasion to refer to this subject were it not that syndactyly is very constantly associated with the split-foot deformity, both in single and familial cases. Further syndactyly has been frequently attributed to this cause. Cleft or deficient palate has also been met with in two split-foot cases. The question can however be dismissed in a few lines. Were the theory of the production of syndactyly in this way tenable, it would raise the needless question of a duplex origin for the combined deformity in the majority of cases, for split-foot can certainly not be accounted for in like fashion. But there is evidence that so arise few, if any, cases of "webbed" fingers. This has been shown by Sutton, Windle and others, and the arguments against it apply with particular force to the cases where the syndactyly is incomplete or bony.

(*e*) *Atavism.* This is often mentioned in connection with split-foot deformities, but in such an indefinite fashion that it is impossible to understand clearly the extent and nature of the reversion to which writers allude. There is a suspicion that they have had in their minds forms phylogenetically remote†. And in a few instances[5] the Quadrumana have been tentatively referred to. A deliberate statement on the subject no one has been bold enough to make, and not the slightest trace of evidence has been put forward in favour of the suggestion. It is of chief interest in that polydactyly has received a similar explanation. The two deformities are intimately associated, and as has been stated in the last paragraph a duplex explanation of origin is objectionable. The view that polydactyly may be atavistic seems to be based largely on the observations and deductions of Bardeleben, who investigated the comparative anatomy of the digits. He is answered by Sutton, Windle and others, who attribute polydactyly to a "sport"; and also by Wiedersheim, who concludes that from the palaeontological point of view the condition of "hyperdactyly" loses its supposed atavistic significance.

The ultimate origin of split-foot conditions and their associated defects in man by a "throw back" can therefore be denied, but there still remains the possibility of its transmission by this means. It might be argued that the original deformed members have arisen from a common ancestor, and that the malformation has lain buried in the intermediate unaffected members of the family. Now an affected individual is never to be found arising from a normal parent, himself the offspring

* Kümmel's case is one of split-hand only. It is discussed by this author from this point of view. His views on the origin of such deformities will be subsequently mentioned.

† "Atavistic parts do not belong to forms palaeontologically remote or systematically far distant" (Sutton, p. 135).

of a deformed individual. Yet in one case a member of a deformed family is said to have had syndactyly only (Mayer), and in another (Anders) the grandfather was reported as having syndactyly of the hands. But from the accounts it would seem that one at least of these cases was not examined, and that consequently foot affection cannot be excluded, and in the other similarly there is no mention of feet*. Thus while a possibility remains it can nevertheless be stated as improbable in the highest degree; the records contain cases in the Mongolian races[12, 13, 18]; also the absolute identity of the defect in different cases is absent. There is no example of intermarriage of two undeformed members of deformed families, an instance of which might throw light on the matter.

Lastly atavism might be brought forward in explanation of the fact that a variety of split-foot deformity may skip one generation to reappear in the next. It remains to be shown however that the deformity in grandfather and child is identical as well as similar.

(*f*) *Origin as a result of Intrauterine conditions.* The question of the origin of split-hand or split-foot by reason of intrauterine lesions is of frequent occurrence in the literature, and the German authors have devoted special attention to it. This explanation has been accepted by Kümmel for all such malformations, and for some of them by others[13, 22]. It has been supposed that the deformities may be produced in the following ways.

1. By an injury to the abdomen from without.

2. By an amputation of digits through strangulation.

3. By the pulling away of the central digit during the contraction of adhesions or the growth of the child.

4. By the splitting of hands or feet through the pressure of bands or folds of amnion, or by the umbilical cord.

The majority of writers have been dealing with single cases, and so far as single split-hand and atypical single split-foot are concerned it cannot be denied that this explanation will in many cases suffice. Kümmel has been criticised by Perthes for extending his hypothesis to hereditary cases, but the arguments of the last author lose their strength if there is any possibility of the transmission of characters acquired *in utero*. We are unaware of a case of single split-foot which can be brought within the category of the cases we are discussing, and great difficulties in accepting the hypothesis are met when the symmetrical and hereditary cases are dealt with.

Intrauterine conditions might be involved in the production of deformities in

(1) the original deformed individuals, whose deformities are transmitted;

(2) the original individuals whose deformities are not transmitted;

and (3) in those individuals to whom the deformity has been transmitted.

* The only reference to the grandfather in Anders' paper is as follows :—"Der Grossvater des Kindes litt an Syndactylie einer Hand."

Lastly they might be involved at a very early or at a later period of development.

The views, which have been held, deal, as may be seen from what has already been said, with the origin at a later date. So far as they apply to the origin of the malformation in those members of a family who have inherited the defect*, they may be at once excluded, for the lesions are in the four extremities in the majority of cases; they are fairly symmetrical; they are similar in different individuals; there is no evidence of similar and simultaneous affection of any other part† of the body; and one of two twins may be alone affected[4, 9, 11]. Again actual evidence of bands, adhesions, or amputated digits has so far been in no case forthcoming‡. The argument that there should be transmission through the female only is insufficient, for the amnion certainly and perhaps the amniotic fluid too is of foetal origin. Lastly the theory involves the supposition that acquired characters are transmitted.

Possibilities 1 and 2 may be considered together, for arguments have been already produced to show that all such individuals are examples of an entity, and that all such deformities would be transmitted if the opportunity arose§. Now if these original members are examined it is found that they differ in no way from the later members of families. As a whole they show the same general characters and the same variations. And this similarity suggests that a similar fundamental factor is at fault in the production of the original and subsequent deformities. But as has been previously stated the mode of origin is not necessarily identical, and it can only be stated that in all probability there is much in common in their production. Further arguments have consequently to be considered. The evidence which refutes the possibility of the origin of deformities in the later members of a family in intrauterine lesions applies to a great extent to the original members, particularly the symmetry, the quadruple nature of the deformity, the similarity of the separate individuals and the absence of evidence of such lesions. And these facts are stongly in favour of the existence of the cause, whatever it may be, at a time when a common factor may act simultaneously upon four masses of formative material, representing the four extremities respectively, or of its existence at a time when all four extremities have a common representative. But in any case the evidence necessitates the assumption that the cause exists at a very early period of development, and certainly prior to the division of the ovum. The great functional capacity of the deformed limbs is in itself a proof of the early laying down of the hands and feet in their malformed state.

* It is to these that the arguments of Perthes particularly apply.

† It must be noted that in two single cases the deformity has been associated with deficiency of the palate.

‡ In the case of J. A. (V, 38) which we have described anything of the kind certainly did not exist.

§ Weismann criticising Zander's opinion that polydactyly arises as a result of the action of amniotic threads on the finger-buds says that such a view is untenable in cases where there is a hereditary history.

Abnormal intrauterine conditions can consequently be rationally assigned as the cause of the deformity only if it be supposed that they are present at this early period; and if it be granted that the origin of the original deformed and later members takes place by the intervention of similar agencies, a position which is strengthened by considerations already alluded to, it becomes practically certain that the process must be regarded as occurring at a still earlier epoch. For a deformed man may have as in Mayer's family deformed children by separate and normal wives, and therefore the deformity must exist in a potential form in the germs of the father.

In the next section this line of reasoning will be further pursued.

(*g*) *Origin as a sport.* The main alternatives have now been considered and, by a process of exclusion and from the outcome of the discussion upon the relation of intrauterine conditions to deformities, it may be concluded that the factors concerned in the causation of hereditary split-foot exist very early. The deformity is in fact a sport, and its association with polydactyly and syndactyly confirms this conclusion.

Weismann was of opinion that polydactyly is due originally to a germ variation, and considered that it must be so when the deformity is transmitted. Windle states that two main views have been held in the case of duplicity, either that the deformity has arisen from the inclusion of additional germ plasm, or that a fission of formative material has taken place. It might similarly be concluded that conditions of deficiency are due to deficiency of germ plasm, or to a failure of fission of the formative material. So far as ectrodactyly pure and simple is concerned, these explanations might suffice, but in the case of the variety of split-foot which we are discussing they are insufficient. For the deformity may be associated with polydactyly, and with bony syndactyly of such a character that it alters the whole conformation of the hand. Further they fail to offer a rational explanation of the differences found in individual sports, which are in all probability collectively an entity. It is in the general conformation rather than in the development or non-development of individual bones that the fundamental defect lies. The normal regularity of conformation and the morphological resemblance of hands and feet can only be explained by assuming that, in addition to the presence in the germ cells of determining factors for the constituent parts of the extremities, there must exist factors determining the arrangement and growth of these, either separately or collectively (cp. Weismann). There is consequently in the case of certain deformities a necessity for assuming a defect in those factors which govern the conformation of the extremities, when the allocation of the origin of the deformity to a particular stage of development is desired. Theoretically the defect may arise at any period in the life history of either germ cell or in the fertilised ovum prior to its division, or by a combination of certain factors in both germs at the time of union. But the latter supposition necessitates the superfluous assumption of variation in the germ cells of both parents. Also both this and the assumption that the variation may occur in the ovum itself,

which can only be allowed if it can be shown that characters acquired *in utero* may be transmitted, are discountenanced by the statement of Vrolik that a normal father may give rise to two similar sport offspring from two separate normal mothers. There is however no recorded example of this sort as far as split-foot is concerned, and the argument therefore rests upon analogy; but it may be said that on the whole the evidence is in favour of the presence of potential deformity, such as split-foot, *in those factors in the germ cell which eventually influence the conformation of the hands and feet.*

When the enormous waste of sexual cells from both male and female genitors in the process of procreation is considered it becomes extremely probable that a vast number of sperms and a large number of ova contain the necessary generating factors of the deformity; for otherwise the unlikelihood of a malformation ever showing in the next generation would be overwhelmingly great. The hypothesis is much strengthened by the fact that more than one deformed child may arise from the same normal parents[8, 13]. It must also be remembered that at least 30 such sports have been reported as occurring. If the position is accepted it follows that when all sports are considered a large proportion of the human race have the potentiality for producing such defects in their offspring. Also it follows that, if a similar potential deformity exists in an abundance of cells developed from a single individual, we have to choose between the possibilities of the origin of the determining factor in a single common precursor of these cells, and the origin of similar defects in each germ cell separately. The former is obviously the more plausible*. It remains for us to further discuss the nature of the particular sport with which we are dealing, and in so doing shall prepare the way for the consideration of certain suggestions relating to sport characters, which have recently been prominent. It has already been pointed out that there are records of over 30 instances of the origin of this same sport†, and that whenever it occurs it tends to vary, when opportunity is given it, in the different individuals of the families. An apparent example of a sport tending to multiple variations, and one which we are inclined to regard as parallel to hereditary split-foot, is the remarkable plant *Oenothera Lamarkiana* and its mutations, discovered by de Vries. In all cases where families of split-foot have been described the variations have occurred, but what is no less remarkable is that the same sport, arising on so many separate

* In assigning the origin to precursors of the sperm cells or ova, the unusual and recently discovered types of nitosis in these cells are not forgotten. But at the same time we cannot commit ourselves to a theory of origin or transmission through the medium of any particular anatomical structure. It is for this reason that we have avoided the use of the word "determinant," as it is too closely associated with these theories.

† In regard to this statement it is essential to point out that in a small proportion of the cases, there has been no mention of descent from healthy parentage; the possibility of two or more of them belonging to the same family can consequently not be placed entirely out of court, though its improbability is great. Moreover it must be remembered that although in the families examined there is no instance in which the deformity skips a generation, yet the presence of deformity in collateral branches cannot be absolutely excluded. Our subsequent conclusions must therefore be read with due regard to these possibilities.

occasions, has shown variation along lines showing close resemblance. The facts are so striking as to suggest that nature's sports are not so erratic as they might at first sight seem, *but that they tend to occur in definite directions and along definite lines.*

(*h*) *Transmission of hereditary split-foot, its stability, and the relation of its mode of transmission to Mendelism.* The arguments which assign the origin of hereditary split-foot to a sport have been fully considered and we are now in a position to discuss the mechanism of transmission from this point of view. In dealing with the types and combinations of types sufficient has been said to show the similarity of these in different families. In any particular family* it is impossible to predict the exact deformity or even the type of deformity which will occur in the offspring of an abnormal parent; all that can be prophesied is that the deformed children will conform to one or other of the general types. In a fair percentage of cases however direct transmission of a type from parent to offspring or indirect to a grandchild through a child of a different type is found. The nearest approach to uniformity of type is presented by Mayer's family in which split-hand occurs in almost all the individuals reported; in one instance through three generations. In the same family is an example of a father with both feet and the right hand affected transmitting the same defects to his son (Fig. 2, II, 2 III, 10), also an instance of transmission from father to son of deformity confined to the feet (Fig. 2, II, 3, and III, 11). Two examples of the last are also to be found in Parker and Robinson's tree. In Fotherby's report polydactyly occurs frequently and is transmitted once. In our own family polydactyly occurs in grandfather, mother and daughter (IV, 3; V, 6; VI, 1). Transmission missing one generation is illustrated in Parker and Robinson's family where three grandchildren "revert" to the type of the grandparent who was deformed in all extremities, while the mother of the children had perfect hands. The same occurs once in the "G" family (II, 2; III, 7; IV, 31). The limitation of a type to a branch of a family is fairly well seen in the occurrence of cross-bones in all the children of IV, 23. But although these instances occur with sufficient frequency to merit attention, yet there is no constancy of the inheritance of individual peculiarities.

As these examples are the only striking instances of particular transmission happening in the four fully reported families mentioned, it may be said that they are the exception rather than the rule. And from the instances cited it follows that though examples of direct transmission of types occur too frequently to allow of their explanation as coincidental, yet there is little evidence of their transmission in a stable form†, and it must be admitted that while a lesser deformity (such as that of feet only), may occur, nevertheless the individual *showing* this lesser malformation may transmit the full defect to offspring. The actual deformity in an individual is thus no index of the potential deformity transmitted to that individual.

* In the following, the larger families "G," 5, 9, and 11, are alone dealt with.

† It is true that examples of split-foot and polydactyly of the hands are reported in three successive generations, yet in no case are the deformities identical.

The partial suppression of deformity in one generation has been referred to in dealing with atavism, but the latter, even if applicable as a term, offers no reasonable explanation of the phenomenon. It might be supposed that this partial "latency" could be brought about by the intervention of other characteristics which might themselves be transmitted*. As it is only in exceptional cases that the deformity of a parent is a gauge, and even then a by no means strictly accurate one, of the malformation which will appear in the offspring; and as at the same time the deformities in both parent and offspring, as in all members of a family, may be said to have been cast in the same mould, there is strong presumptive evidence that the inheritance depends on the transmission of some common factor. That representatives of the deformity are not transmitted in detail is obvious, for no two individuals have as yet been reported in any family in which the deformities were identical. Also in view of these multiple variations, it appears inconceivable that the originating factor is a recurring variation in the representatives of the hand or foot in the germ cells. It seems far more probable that a fundamental factor, which influences the ultimate general conformation of the affected parts through their normal representatives, is at fault, that it is transmitted, and that its interaction with these representatives varies slightly in quality and quantity in different individuals and generations; and that the varying interaction is produced by the interference of factors which may or may not be transmitted, such as those which may be conceived to account for the partial latency above mentioned. Such an hypothesis would not only offer a rational explanation of partial latency, but would render the quadruplicity of this and many other deformities less inexplicable; at the same time accounting for the marked tendency to symmetry which usually exists. The hypothesis is in accord with the observations of Weismann on the subject of polydactylism (cp. Wilson and Windle also). Whatever view be taken it is difficult to avoid the conclusion that in spite of the variation in individual deformities, there is a common factor and fundamental scheme in transmission, both as far as hands and feet are concerned.

Before passing to the inheritance of the deformity as a whole there remain two considerations; the first in connection with sex preponderance, the second with twins.

In Bédart's short family there was a predominance of the female sex in the proportion of 2 to 1, though taking deformed and undeformed together the male element was in the ascendant. In Mayer's report there is a male excess in the proportion 6 to 1 (Fig. 2). Apart from these examples sex plays no part in transmission, and therefore requires little comment.

There have been in all four cases of twins. In two of these[4, 11], where the twins were heterologous, the female was in each unaffected. In a third instance (V, 33, 34), two children of opposite sex were normal. Lastly of Mayer's homologous twins (Fig. 2, III, 6, 7), arising from a deformed father, one only was deformed. Interest attaches itself to this case in view of the opinion that certain

* A phenomenon related to if not identical with the "cryptomerism" of Mendelians.

homologous twins result from the fission of a single ovum. We can only conclude that in the case of Mayer's twins such an origin is highly improbable.

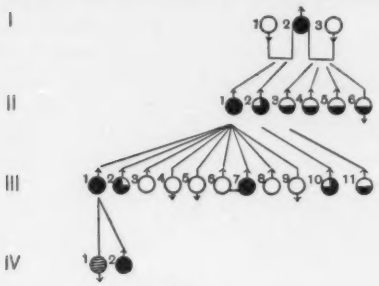



Fig. 2.

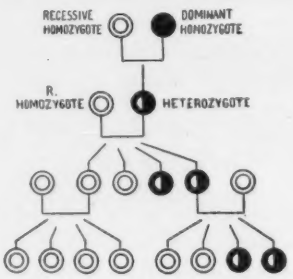



Fig. 3.

Fig. 2. The figure illustrates the deformities of a family of hereditary split-foot described by Mayer (loc. cit.). The arrangement and signs correspond to those of Fig. 1. The family is of special interest as it shows:—marked dominance of the deformity: offspring from two wives by the original deformed member: male predominance: inheritance of type (II, 2 and 3, III, 10 and 11): and homologous twins of which one is deformed (III, 6 and 7).

Fig. 3. Diagram illustrating Mendelism as applied to human sports. The horizontal and oblique lines are to be interpreted as in Figs. 1 and 2. Each individual is represented by a double circle, the outer representing the appearance of the individual (black deformed and white normal), the inner circle divided by a vertical line the nature of the gametes. Homozygotes are consequently represented uniformly in black or white; heterozygotes with the outer ring black (signifying the actual deformity and illustrating the dominance of the same) and the inner ring half black and half white, to denote the potential deformities of the gametic cells. In the figure a recessive and dominant homozygote are supposed to be crossed, producing heterozygotes only, one of which is shown. This individual is then crossed with a normal or recessive homozygote with the production of equal numbers of homozygotes and heterozygotes. Of these offspring one of each variety is crossed with a normal homozygote. The figure shows that from heterozygotes when crossed with normal individuals an equal number of deformed and normal offspring are to be expected.

We now come to a subject of considerable interest and importance, in the discussion of the transmission of the deformity as a whole. There appears to have been of late years a growing feeling that slow or continuous variation is insufficient to account for the origin of widely differing species (cp. Lock), and a tendency to attribute such species to discontinuous variation or sport. The essential consideration in view of the possibility of a new species originating in a sport, is the stability of the particular deformity in its transmission. Weismann stated that there is no difficulty in understanding the gradual swamping of sports, when it occurs, and even went so far as to assert that it is inevitable in the absence of inbreeding. But Weismann based his statements chiefly on theoretical considerations, and the "id" hypothesis in its application to sport characters is perhaps no longer free from doubt. In the recent revival of Mendelism a great many new observations have been made into the subject of "hybridism." The investigations have been made for the most part upon the vegetable and lower

animal kingdom, and the term "hybrid" is now given wider scope. Bateson in a recent communication has suggested the applicability of Mendelian laws to hereditary pathological and teratological conditions. In the ensuing paragraphs the applicability of these laws to hereditary split-foot will be examined; for if the cross between a deformed individual and a normal one is to be designated as a hybrid, and if the transmission of any deformity can be shown to proceed along Mendelian lines, the question of permanent stability of that deformity upon transmission may be considered as settled in the affirmative, and the possibility of a human sport originating a species, in the wide sense of the term, will be past denial.

According to the Mendelian doctrine animals and plants cannot be considered, from the point of view of heredity, as "units," but as composites of separate characters. If as a result of the crossing of two individuals, bearing separate types of a single character, these types segregate independently in the offspring, they are termed "allelomorphs" and may be expected to obey certain fixed laws. A fundamental proposition of Mendelism states that it is impossible for the same gamete to carry more than one member of the same pair of allelomorphs. Again, one allelomorphic character may show dominance to the other member of the pair, in which case the result of the cross between a "dominant" and "recessive" allelomorph will be the appearance of offspring all bearing the dominant character*. The offspring will nevertheless produce an equal number of gametes carrying dominant and recessive characters. Such offspring are termed "heterozygotes," as opposed to those arising from the union of like allelomorphic gametes, which are termed "homozygotes." When a heterozygote is produced by the union of a dominant and a recessive allelomorph, and is crossed with a homozygote resulting from the union of two recessive allelomorphs, homozygous offspring bearing the recessive character and heterozygotes bearing the dominant character will be produced in equal numbers. This principle is illustrated in the accompanying diagram (Fig. 3).

From this necessarily short account† it may be gathered that when dealing with allelomorphic characters an abnormal offspring cannot arise from recessive homozygotes, and that in a family tree arising from an abnormal ancestor the deformity may be expected to be stable. It also follows that the abnormal offspring of one normal and one abnormal parent is a heterozygote.

We have seen that there is reason to believe that a sport in all probability arises as the result of the union of a normal and an abnormal gamete. It may be therefore assumed, if Mendelism is to apply, that the original sport is a heterozygote, an assumption which Bateson appears to have made. From this it follows that if a family is in agreement with Mendelian laws the total number of deformed and undeformed offspring from deformed parents should be equal.

* If the sport character is recessive, it can never appear in the family unless there is intermarriage.

† Which for simplicity takes no account of "latency."

The counts for split-foot families are as follows:—

Parker and Robinson's...16 deformed and 16 normal.
Ours .....44 deformed and 32 normal.
Fotherby's.....16 deformed and 10 + normal.
Mayer's .....12 deformed and 6 normal.

Thus there is a greater dominance* of the deformed than is to be expected, though several of the families are fairly in agreement with the rule. But we have now to proceed a little further.

It has been stated that the original sport should be a heterozygote, but in the family of Mayer (Fig. 2), all the members of the second generation show the dominant character, namely are deformed. On Mendelian principles this would be brought about if the original sport were a homozygote. But if this is assumed there is every reason to assume that every original sport of the sort is a dominant homozygote also, and this cannot be allowed, as were it so all children of the second generations should be invariably deformed, which they are not. Moreover it is the rule in these families for an excess of deformed individuals to show itself in the first generation, and this is the case also with the majority of deformities of the hands or feet which run through long families†.

Lastly if Mendelian proportions are to hold good they must not only hold good in every family of the particular deformity examined, *but they must above all hold good for each succeeding generation*, for upon this depends entirely the question of stability, which is the chief point at issue. In families showing split-foot the following figures are shown:—

| | Parker & Robinson's | | Ours | | Fotherby's | | Mayer's | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | D. | N. | D. | N. | D. | N. | D. | N. | D. | N. |
| 1st gen.‡ ... | 1 | 6 | 1 | 0 | 1? | ? | 1? | ? | — | — |
| 2nd gen. ... | 9 | 4 | 11 | 3 | 1? | ? | 6 | 0 | 27 | 7 |
| 3rd gen. ... | 7 | 12 | 13 | 5 | 2 | 2+ | 5 | 6 | 27 | 25 |
| 4th gen. ... | — | — | 17 | 22 | 9 | 3 | 1 | 0 | 27 | 25 |
| 5th gen. ... | — | — | 1 | 2 | 4 | 5 | — | — | 5 | 7 |
| | | | | | | | Totals ... | | 86 | 64 |

D. = deformed, N. = normal.

An examination of these figures shows a decided tendency for the deformity to die out. In view of this and of the hyperdominance of the deformity as a whole,

* A dominance which has already been noted by Bateson in many families.

† Families illustrating "latency," are not here considered.

‡ For convenience sake, the numbering of generations in this table does not correspond with that given in Fig. 1.

the evidence points to the conclusion, despite the apparent segregation, that the transmission of hereditary split-foot is not governed by Mendelian laws. We have not yet seen a family tree of fair proportions which shows this adherence to Mendelism*. The family trees given by Bateson tend to show the same diminution in the proportion of affected to non-affected offspring in succeeding generations. In concluding this paragraph it may be said that evidence has not yet been produced which warrants the statement that any human sport is transmitted along the lines of Mendelism.

Passing back to the original question, we find that the evidence of stability which Mendelism might have given is not forthcoming and that when the figures are taken on their own merit, they uphold the view expressed by Weismann that swamping of transmitted sports will eventually take place. As Weismann stated for polydactyly, that no family deformity has been traced through six generations, so might the same statement be made for hereditary split-foot, but the statement of Weismann no longer holds good† and there is no reason to disbelieve that the deformity of the "G" family will assert itself for several if not many more generations.

Given the factor of inbreeding a split-foot race might quickly arise. In this connection the family quoted by Windle after Devay is of interest (*Linnean Soc. Trans.*), for it shows the development of a polydactylous race in an isolated village.

Deformities may die out in one of several main ways. The deformity may be so gross as to render life impossible. Lesser grades of malformation though compatible with extrauterine life may predispose to early death, or handicap the individual in obtaining the necessaries of life. Other deformities, from their tendency to disfigure, may hinder mating. Others may render procreation impossible. Of these none, so far as can be seen, influence to any extent the prolificacy of individuals suffering from hereditary defects of the digits. In Devay's family there appears, upon the introduction of new blood, to have been a gradual diminution in the severity of the individual lesions as they passed through succeeding generations. The same tendency was considered by Fotherby and Mayer to be present in their families, but the evidence in these last cases is inconclusive. The extermination of hereditary split-foot takes place by a proportionate decrease in the number of deformed offspring, arising from deformed parents, from one generation to the next. To what this diminution is due is a question to which at present no answer is forthcoming, but the fact as it stands is directly opposed to the supposition that human sports or their originating factors may be regarded as unit characters, transmitted along those definite lines which certain natural characters in animals have been shown to be inherited.

* Though we have searched a large number.

† Families of eight generations have been reported.

*Summary and Chief Conclusions.*

1. A family containing 44 deformed members is placed on record. Of these 17 are described in detail.

2. Of the many varieties of split-hand and split-foot, one is most prominent and is of sufficiently common occurrence to merit the name "hereditary split-foot." A general account of it is given and reasons assigned for considering it a deformity *sui generis*.

3. It usually affects the four extremities and shows great variability in all families in which it occurs. The types of deformity of the hands and feet are discussed and their relationship shown in different individuals and in separate families.

4. All the cases of symmetrical split-foot, which have been found, have been included in the bibliography. Of these over 167 occur in families in which other members are similarly affected; the remaining 13 are isolated. Nevertheless there is reason to believe that they are all of the same nature.

5. The deformity consists mainly of an ectrodactyly, but is frequently associated with various grades of syndactyly and polydactyly.

6. The deformity has its origin in a "sport," which takes place in the parental germ cells or their precursors, probably in the latter.

7. A plausible hypothesis is advanced in explanation of the quadruple nature of the sport and its tendency to recurring variation. It is supposed that the mutation originally occurs in that factor in the gametic cell (or its precursor) of the parent which governs the general and eventual conformation of the hands and feet, and that this mutation is transmitted as such. The succession of split-foot members in a family is consequently regarded as due to the transmission of a tendency to sport along definite lines.

8. It is probable that nature's sports occur in definite directions.

9. The transmission of "hereditary split-foot" does not follow the laws of Mendel. It is true that the deformity segregates, but it appears to a diminishing extent in succeeding generations. It shows no tendency to skip generations.

10. It remains to be shown that this or any other human sport is permanently stable, or strictly follows Mendelian laws.

## BIBLIOGRAPHY.

### A. *"Hereditary split-foot" families.*

1. Anders (E.). Jahrb. f. Kinderheilk. 1880–81, Bd. 16, S. 435. (3 cases.)
2. Anderson (W.). B. M. J., June 12, 1886, p. 1107. (24 cases.)
3. Béchet. Essai, Paris, 1829 (cited by Hilaire, and Fort, loc. cit.). (5 cases.)
4. Bédart. Rev. Mens. de l'École d'Anth. de Paris, année 2, p. 244 (cited by Windle, loc. cit.) (12 cases.)
5. Fotherby (H. A.). B. M. J., 1886, May 22, p. 975. (17 cases.)
6. Goldman (E. E.). Beiträge z. klin. Chir. 1891, Bd. 7, S. 249. (3 cases.)

7. HOLMES (T.). "Surgical treatise of diseases of infancy, etc." Lond. 1868, p. 215; and "Thérapeutique des Malad. Chirurg." Paris, 1870, p. 322. (4+ cases.)
8. JAYLE et JARVIS. Bull. de la Soc. Anat. de Paris, 1898, p. 139. (3 cases.)
9. MAYER (C.). Beiträge z. path. Anat. (Ziegler), 1898, Bd. 23, S. 20. (14 cases.)
10. MAYGRIER. Inaug. Dissert. (cited by Gould and Pyle, loc. cit.) (4 cases.)
11. PARKER (R. W.) & ROBINSON (H. B.). Clin. Soc. Trans. 1887, Vol. 20, p. 181. (17 cases.)
12. PASTER (Cl.). Virch. Archiv, Bd. 104, S. 54. (2 cases.)
13. PERTHES. Deutsch. Zeitsch. f. Chirurg. 1902, Bd. 63, S. 132. (5 cases.)
14. POTT (R.). Jahrb. f. Kinderheilk. 1884, Bd. 21, S. 392. (4 cases.)
15. SCHÄFER (W.). Beiträge z. klin. Chir. 1892, Bd. 8, S. 436. (9 cases.)
16. SCOUTETTEN (H.). Bull. de l'Acad. Impér. de Méd. 1857, T. 23, p. 97. (7 cases.) (Contemporary account by Fort, gives 9 cases.)
17. TILANUS. Zeitsch. f. orth. Chir. 1896, S. 398 (cited by Windle, loc. cit.) (3+ cases.)
18. THOMSON (J. D.). B. M. J., 1892, Vol. 1, p. 1188. (4 cases.)
19. TUBBY (A. H.). Lancet, 1894, Vol. 1, p. 396; and "Deformities." Lond. 1896, p. 506. (15 cases.)

### B. *Single cases of "hereditary split-foot."*

20. COATS (J.). "Manual of Pathol." London, 1903, p. 60. We are indebted to Dr Sutherland, the present editor, for notes of the case.
21. Ewh. Inaug. Dissert. Witten, 1890 (cited by Kümmel, loc. cit.).
22. FRONHÖFER (E.). Archiv f. klin. Chir. (Langenbeck), 1896, Bd. 52.
23. GAILLARD. Gazette Médicale de Paris, 1859, T. 40, p. 787.
24. GOULD (A. P.) & PYLE (W. L.). "Anomalies and Curiosities, etc." London, 1901.
25. GUERMONPREZ (Fr.). Bull. et mém. de la soc. de chir. 1884, T. 10, p. 722.
26. KING (C.). Internat. Med. Mag. Philadelph. 1897, Vol. 5, p. 91.
27. MOREL-LAVALLÉE. Bull. et mém. de la soc. de chir. 1884, T. 10, p. 726 (extr.).
28. MELLER (J.). Berlin klin. Wochensch. Marz 6, 1893, S. 232.
29. MENIÈRE (P.). Archiv. gén. de Méd. Paris, 1828, T. 16, p. 364. (2 separate cases.)
30. OTTO (A. W.). "Monstrorum Sexcentorum Descriptio Anatomica." Vratislaviae, 1841, p. 125, etc., Tab. 17–20.
31. PYE. Lancet, 1889, Vol. 2, p. 1119.
32. WIPPER. Inaug. Dissert. München, 1886 (cited by Kümmel, loc. cit.).

### C. *Cases of split-hand, with feet unaffected.*

33. ANNANDALE (T.). "Diseases of Fingers and Toes." Edinburgh, 1865.
34. DOWD (C. N.). Annals of Surgery, 1896, Vol. 24, p. 211.
35. FUMAGALLI. Schmidt's Jahrbuch, 1872, Bd. 153, S. 136 (review by Thiele).
36. GIRALDES (J. A.). "Leçons clin. sur les Malad. Chir. des Enfants." Paris, 1869.
37. GUERMONPREZ (Fr.). Loc. cit. supra.
38. KORMANN (E.). Jahrbuch f. Kinderheilk. 1880, S. 420.
39. KÜMMEL (W.). Bibliotheca Medica, Abth. E, Hft. 3, 1895.
40. PAULICKY. Deutsch Militär. Zeitsch. 1882, Bd. 11, S. 199 (cited by Kümmel, loc. cit.).
41. SCHÄFER (W.). Loc. cit. supra.
42. WALTHER (P.). Journ. d. Chir. u. Augenhk. 1829, Bd. 13, S. 373 und Taf. 6.

### D. *General References.*

43. AMMON (F. A.). "Die angebor. chirurg. Krankh. d. Menschen." Berlin, 1842, Bd. 4, Abth. 1.
44. BARDELEBEN (K.). "Hand und Fuss." Verhdl. d. Anat. Ges. Jahrg. 8, 1894, S. 257.

45. Bateson (W.). "Mendel's Principles of Heredity." Cambridge, 1902; and Brain, 1906, Pt. 104, p. 157.
46. v. Bergmann (E.)*. "Handbuch d. pract. Chir." Bd. 4, Abth. 1, S. 339, and Abth. 2, S. 455.
47. Devay. "Des Dangers des Marriages Consanguins." Paris, 1862, p. 95 (cited by Windle, loc. cit.).
48. Förster (A.)*. "Die Missbildung des Menschen." Jena, 1861.
49. Fort (J. A.). "Des déformités congénitales, etc." Paris, 1869.
50. Gintrovicz. Inaug. Dissert. Würzburg (cited by Kümmel, loc. cit.).
51. Hilaire (Geoffroy Saint). "Histoire gén. et part. des Anomalies de l'organisation, etc." 1832, T. 1, p. 676.
52. Lewis (H. F.). Amer. Journ. Obstet. 1899, Vol. 40, p. 84.
53. Lock (R. H.). "Recent progress in Variation, etc." London, 1906.
54. Murray (J. J.). Med. Chir. Trans. 1863, Vol. 28, p. 29.
55. Sutton (J. B.). "Evolution and disease." London, 1890.
56. Vrolik (G.). Todd's Cyclopaed. Anat. and Physiol. London.
57. Weismann (A.). "The Germ Plasm, etc." (Translated by Parker, etc. London, 1893.)
58. Wiedersheim (R.). "The Structure of Man." (Trans. by Bernard. London, 1895. p. 81.)
59. Wilson (G.). Journ. Anat. and Physiol. 1896, Vol. 30, p. 437.
60. Windle (B.). Journ. Linnean Soc. 1891, Vol. 23, p. 488; and Journ. Anat. and Physiol. 1891–92, p. 295; and ibid. 1893–94, p. 25.
61. Young (J. K.). "A Manual of Orthop. Surg." London, 1906. (This author refers to two families showing lobster-claw hand.)

## EXPLANATION OF PLATES.

Plate I. Fig. 1. Genealogical tree of the "G" family.

The following signs are used :—

The horizontal lines are drawn from parent to parent.
The oblique lines are drawn from parents to offspring.
The wavy lines indicate uncertainty of order of birth.
Black represents deformity, each quadrant of the circle representing the corresponding limb.
The shaded circles are instances of doubtful deformity.
The arrows when present indicate sex.
M. signifies a miscarriage. S.P. = died without offspring. D. = diseased. Illeg. = illegitimate.
Generation III, groups 3 and 4 represent 9 deformed and 2 normal offspring.
In generation V, 33 and 34 are twins and the circles are joined by a single horizontal line.

Plate II. Figs. 4 and 5. Photographs of the hands and feet of R. E. G. (Fig. 1, V, 41). The photographs show typical examples of the deformity in its external configuration.

Plate III. Figs. 6—9. Left hand, right hand, left foot and right foot respectively of R. E. G. (V, 41). Note the symmetry.

Plate IV. Figs. 10 and 11. Left and right hand of M. A. (IV, 17). Figs. 12 and 13. Left and right hand of E. M. G. (V, 40).

Plate V. Fig. 14. Right hand of L. V. G. (V, 45). Fig. 15. Right foot of E. M. G. (V, 40). Fig. 16. Left foot of M. A. (IV, 17). Fig. 17. Left hand of W. H. A. (V, 24); note that all skiagrams with this exception are taken from the dorsal aspect.

Plate VI. Figs. 18 and 19. Left and right hand of J. T. G. (V, 42). Figs. 20 and 21. Left and right foot of L. V. G. (V, 45).

Plate VII. Figs. 22 and 23. Left and right hand of E. G. (IV, 23). Figs. 24 and 25. Left and right foot of J. T. G. (V, 42).

* Certain of these authors give isolated cases of split-foot, but they are apparently taken from sources already referred to.

# ON THE GENERALISED PROBABLE ERROR IN MULTIPLE NORMAL CORRELATION.

By KARL PEARSON, F.R.S., and ALICE LEE, D.Sc.

The normal correlation surface in the case of $n$ variables is known to be:

$$z = \frac{N}{(2\pi)^{\frac{1}{2}n}\sigma_1\sigma_2 \dots \sigma_n \sqrt{R}} e^{-\frac{1}{2R}\{S(R_{pp}x_p^2/\sigma_p^2)+2S'(R_{pq}x_px_q/\sigma_p\sigma_q)\}},$$

where

$$R = \begin{vmatrix} 1, & r_{12}, & r_{13}, & \dots & r_{1n} \\ r_{21}, & 1, & r_{23}, & \dots & r_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ r_{n1}, & r_{n2}, & r_{n3}, & \dots & 1 \end{vmatrix}$$

and $R_{pq}$ is the minor corresponding to the constituent $r_{pq}$. $S$ denotes a sum for all values of $p$ from 1 to $n$, and $S'$ a sum for every pair of unlike $p$, $q$'s. $x_p$ is the $p$th variate measured from its mean, $\sigma_p$ its standard deviation and $r_{pq}$ the correlation coefficient of the $p$th and $q$th variates.

Let $$\chi^2 = \frac{1}{R}\{S(R_{pp}x_p^2/\sigma_p^2) + 2S'(R_{pq}x_px_q/\sigma_p\sigma_q)\}.$$

Then $\chi^2 =$ a constant, is the equation to an ellipsoid in $n$-fold space. If transferred to its principal axes, it will take the form

$$\chi^2 = \sum_1^n (X_p^2/\Sigma_p^2).$$

Further: $z dx_1 dx_2 \dots dx_n$ will equal

$$\frac{N}{(2\pi)^{\frac{1}{2}n}} \frac{1}{\Sigma_1\Sigma_2 \dots \Sigma_n} e^{-\frac{1}{2}\chi^2} dX_1 dX_2 \dots dX_n,$$

representing the frequency due to $n$ independent variates.

The ellipsoid in $n$-fold space can be reduced to a sphere by proper squeezes in the directions of its principal axes, and accordingly its volume can be found from that of a sphere in $n$-fold space. The volume of a sphere in $n$-fold space

$$= \frac{2^n a^n (\sqrt{\pi})^n}{1.2.3 \ldots \frac{1}{2}n}, \text{ if } n \text{ be even},$$

$$= \frac{2^{\frac{1}{2}(n+1)} a^n (\sqrt{\pi})^{n-1}}{1.3.5 \ldots n}, \text{ if } n \text{ be odd}.$$

Hence the volume of an ellipsoid in $n$-fold space

$$= \frac{2^n a_1 a_2 a_2 \ldots a_n (\sqrt{\pi})^n}{1.2.3 \ldots \frac{1}{2}n}, \text{ if } n \text{ be even},$$

$$= \frac{2^{\frac{1}{2}(n+1)} a_1 a_2 a_3 \ldots a_n (\sqrt{\pi})^{n-1}}{1.3.5 \ldots n}, \text{ if } n \text{ be odd}.$$

Applying these to our special cases we have for the volume of correlation ellipsoid, $V$:

$$V = \frac{2^n \Sigma_1 \Sigma_2 \ldots \Sigma_n \chi^n (\sqrt{\pi})^n}{1.2.3 \ldots \frac{1}{2}n}, \text{ if } n \text{ be even},$$

$$= \frac{2^{\frac{1}{2}(n+1)} \Sigma_1 \Sigma_2 \ldots \Sigma_n \chi^n (\sqrt{\pi})^{n-1}}{1.3.5 \ldots n}, \text{ if } n \text{ be odd}.$$

Thus for the two cases:

$$\delta V = \frac{2^n \Sigma_1 \Sigma_2 \ldots \Sigma_n \, n\chi^{n-1} \delta\chi (\sqrt{\pi})^n}{1.2.3 \ldots \frac{1}{2}n} \quad (n \text{ even}),$$

and

$$= \frac{2^{\frac{1}{2}(n+1)} \Sigma_1 \Sigma_2 \ldots \Sigma_n \, n\chi^{n-1} \delta\chi (\sqrt{\pi})^{n-1}}{1.3.5 \ldots n} \quad (n \text{ odd}).$$

Now the total amount of frequency between two contour ellipsoids is $z\delta V$, and accordingly the frequency inside a given contour ellipsoid $\chi$ is given by:

$$I_\chi = \int_0^\chi z\,dV = \frac{N}{2.4.6 \ldots n-2} \int_0^\chi e^{-\frac{1}{2}\chi^2} \chi^{n-1} d\chi, \text{ for } n \text{ even},$$

$$= \sqrt{\frac{2}{\pi}} \frac{N}{1.3.5 \ldots n-2} \int_0^\chi e^{-\frac{1}{2}\chi^2} \chi^{n-1} d\chi, \text{ for } n \text{ odd}.$$

Let

$$\mu_n(x) = \frac{1}{\sqrt{2\pi}} \int_0^x x^n e^{-\frac{1}{2}x^2} dx$$

be defined as the "incomplete $n$th normal moment function." Then

$$I_\chi/N = \frac{\sqrt{2\pi}\,\mu_{n-1}(\chi)}{2.4.6 \ldots n-2}, \text{ if } n \text{ be even},$$

$$= \frac{2\mu_{n-1}(\chi)}{1.3.5 \ldots n-2}, \text{ if } n \text{ be odd}.$$

Now $I_\chi/N$ is the chance of an observation falling inside the contour $\chi$, and $1 - I_\chi/N$ the chance of its falling in the fringe outside this contour.

We see accordingly that the chance of an outlying observation being reasonable or not,—the observation consisting of a complex of $n$ variates,—can be readily found, if the incomplete normal moment functions have once been tabled, and the constants of the correlation surface be known. These incomplete normal moment functions serve a variety of purposes which will be developed in later papers. The present paper merely refers to the means they provide of determining the probability of any observation lying outside a given contour ellipsoid, i.e. there being any "fringe" beyond this value of $\chi$.

In Table I. pp. 66, 67 the values of $m_n(x)$

$$= \mu_n(x)/\{(n-1)(n-3)(n-5)\dots 1\}, \text{ if } n \text{ be even,}$$

and $$= \mu_n(x)/\{(n-1)(n-3)(n-5)\dots 2\}, \text{ if } n \text{ be odd,}$$

are tabled for values of $x$ from 0 to 5 proceeding by tenths. They were calculated in the following manner. Let

$$q_n = \sqrt{\frac{2}{\pi}}\int_x^\infty e^{-\frac{1}{2}x^2}dx + \sqrt{\frac{2}{\pi}}e^{-\frac{1}{2}x^2}\left\{\frac{x}{1} + \frac{x^3}{1.3} + \frac{x^5}{1.3.5} + \dots + \frac{x^{n-1}}{1.3.5\dots n-1}\right\},$$

if $n$ be even,

$$q_n' = e^{-\frac{1}{2}x^2}\left\{1 + \frac{x^2}{2} + \frac{x^4}{2.4} + \dots + \frac{x^{n-1}}{2.4.6\dots n-1}\right\}, \text{ if } n \text{ be odd.}$$

Then: $$m_n = \tfrac{1}{2}(1-q_n), \text{ if } n \text{ be even,}$$

$$= \sqrt{\frac{1}{2\pi}}(1-q_n'), \text{ if } n \text{ be odd.}$$

The even series involve the probability integral which was taken from Sheppard's Tables (*Biometrika*, Vol. I. p. 182). Consequently the even series may have an error of unity in the seventh figure. It is hoped that arithmetical blunders have been avoided, but the labour has been very considerable and some blunders may have escaped notice.

If we make $I_\chi/N = \frac{1}{2}$, we find the contour ellipsoid within which half the frequency lies. In other words an observation is as likely to lie inside as outside this contour ellipsoid. The corresponding value of $\chi$ may be spoken of as the "generalised probable error." Since the first 10 incomplete moment functions have been calculated it is possible to determine the generalised probable errors for 1 up to 11 variables. These are given in the table on the next page.

Of these, the first is very familiar. The second was first given in Bravais' classical memoir (*Mém. prés. par divers Savans Étrangers*, T. IX. 1846, pp. 255—332). The third professes to be given by Czuber, *Theorie der Beobachtungsfehler*, S. 404, but I do not agree with his results. The value 1·53817 was given by me in College Lectures in 1896. The remainder have, as far as I know, not hitherto been published, although they have been for a considerable time in use in my Biometrical Laboratory.

*Table of Generalised Probable Errors.*

| Number of Variables | Probable Error |
|---|---|
| *1* | 0·674,4898 |
| *2* | 1·177,4062 |
| *3* | 1·538,1667 |
| *4* | 1·832,1239 |
| *5* | 2·086,0146 |
| *6* | 2·312,5982 |
| *7* | 2·519,0869 |
| *8* | 2·710,0022 |
| *9* | 2·888,3962 |
| *10* | 3·056,4366 |
| *11* | 3·215,7402 |

The method of calculation was as follows: If $n$ be even we have from the equations on p. 61:

$$m_{n-1} = \tfrac{1}{2}\frac{1}{\sqrt{2\pi}} = \cdot 1994711.$$

If $n$ be odd: $m_{n-1} = \tfrac{1}{4} = \cdot 25,$

since $I_\chi/N = \tfrac{1}{2}$.

The problem therefore resolves itself into an interpolation in the tables. Let $y$ be the required argument of the table and $x$ the corresponding value of $m_{n-1}$; let $x_0$, $x_1$, be the two values in the table immediately below $x$, and $x_2$, $x_3$ immediately above $x$; then the formula

$$y = y_3 + \tfrac{1}{10}\left\{ -\frac{2(x-x_1)(x-x_2)(x-x_3)}{(x_0-x_1)(x_0-x_2)(x_0-x_3)} - \frac{(x-x_0)(x-x_2)(x-x_3)}{(x_1-x_0)(x_1-x_2)(x_1-x_3)} - \frac{(x-x_0)(x-x_1)(x-x_3)}{(x_2-x_0)(x_2-x_1)(x_2-x_3)} \right\}$$

was used to find $y$, i.e. a cubical parabola was passed through the four points. The accuracy of a variety of other interpolation formulae was found to be less.

Having found the values of the first eleven probable errors, an empirical formula was then sought which would closely express them, and which accordingly might be used for extrapolation. The form of curve finally adopted was that found by adding the ordinates of a line to that of a logarithmic curve. Its equation, after the adjustment of its constants by the method of least squares, was:

$$\text{P.E.} = -\cdot 067,769 + \cdot 071,986n + 2\cdot 308,272 \log_{10}(\cdot 994,484 + n),$$

giving the probable error in terms of the number of correlated variables $n$.

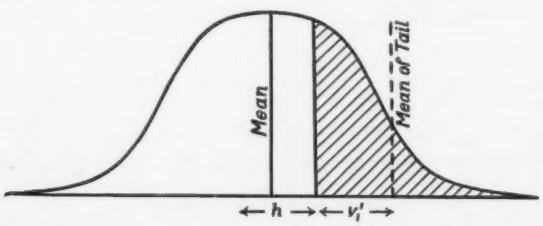
Of course we must not expect such an arbitrary formula to be true to more than some four figures. As a matter of fact in fitting it only the values from 3 to 11 were taken. The following give the calculated and actual values:

| $n$ | Calculated* | Actual |
|---|---|---|
| *1* | ·696 | ·674 |
| *2* | 1·176 | 1·177 |
| *3* | 1·537 | 1·538 |
| *4* | 1·832 | 1·832 |
| *5* | 2·087 | 2·086 |
| *6* | 2·314 | 2·313 |
| *7* | 2·520 | 2·519 |
| *8* | 2·710 | 2·710 |
| *9* | 2·888 | 2·888 |
| *10* | 3·055 | 3·056 |
| *11* | 3·215 | 3·216 |
| *12* | 3·367 | — |
| *13* | 3·513 | — |
| *14* | 3·654 | — |
| *15* | 3·791 | — |

It appears accordingly probable that our formula up to possibly $n=15$ or more will represent the probable error with not more than ·001 error. The extreme divergence for the case of $n=1$ is noteworthy, and seems to place this result on a different footing to the others. The general agreement is shown graphically in the diagram on p. 64.

The general use of the table provided will be obvious, it enables us to tell the probability of any outlying individual really being a member of a population of which the constants are known. Thus one may look forward to the day when the biometric constants of a race being sufficiently well known, it may be possible to tell from a complex of five or six characters whether a skeleton or a skull may be reasonably supposed to have belonged to a member of that race. At present the labour of calculating the correlation coefficient-determinants and their minors stands in the way of much work in this direction, when we wish to advance beyond two or three characters.

The incomplete normal moment functions provide us with a method of determining, theoretically at least, the constants of a truncated normal distribution. The method is as follows. Let $N$ be the total frequency, $\sigma$ the standard




* For plotting the curve we have also the ordinates ·3710, ·0276, – ·0733 for $n=\frac{1}{2}$, 0·1, 0, respectively.

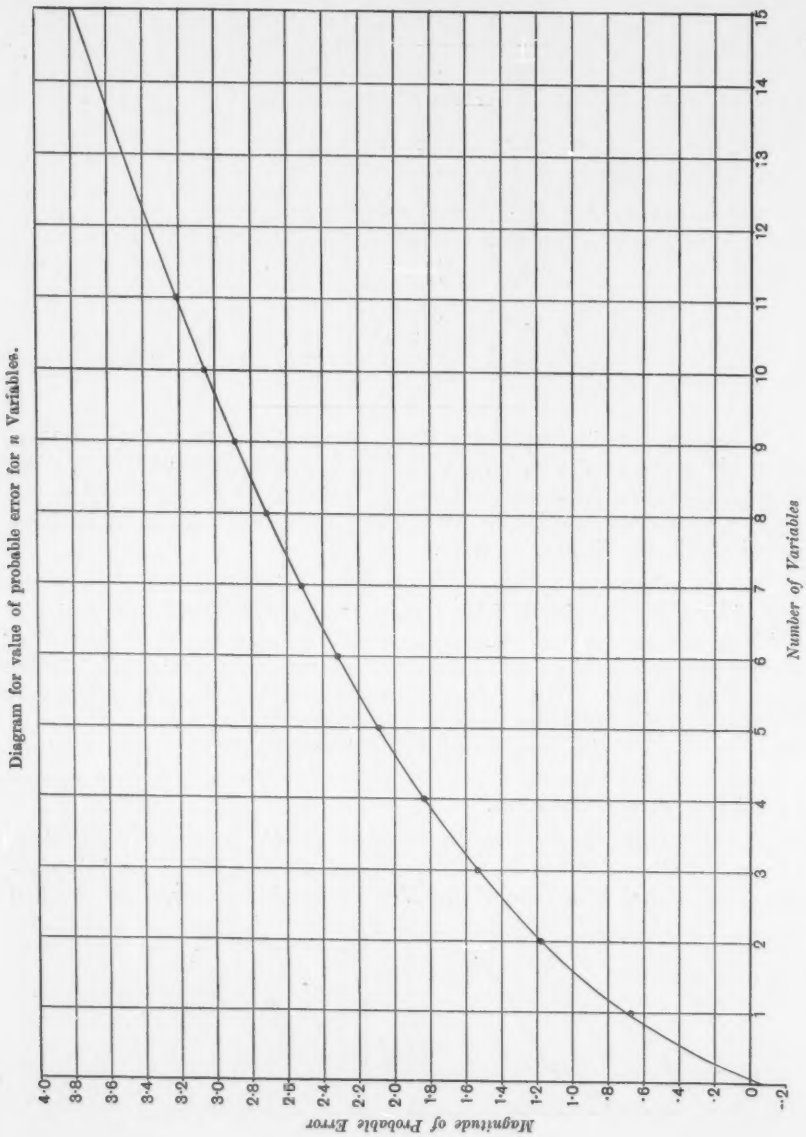


Diagram for value of probable error for n Variables.
Magnitude of Probable Error
4·0
3·8
3·6
3·4
3·2
3·0
2·8
2·6
2·4
2·2
2·0
1·8
1·6
1·4
1·2
1·0
·8
·6
·4
·2
0
-·2
Number of Variables
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15

deviation, and $h$ the distance from the mean at which the distribution is truncated. Then the frequencies of the shaded part are known, but $N$, $h$ and $\sigma$ are unknown. Let $n$ be the known total frequency of the tail, $\nu_1'$ the distance of its centroid, or mean, from the 'stump,' $\nu_2'$ the second moment of the tail about the stump, and $\Sigma$ the standard deviation of the tail, of course about its mean; then $\nu_1'$ and $\nu_2'$ (or $\Sigma = \sqrt{\nu_2' - \nu_1'^2}$), are known from the observations.

Now $n\nu_p'$ is the $p$th moment of the tail about its stump and

$$n\nu_p' = \frac{N}{\sqrt{2\pi}\,\sigma}\int_h^\infty e^{-\frac{1}{2}\frac{x^2}{\sigma^2}}(x-h)^p\,dx$$

$$= \frac{N\sigma^p}{\sqrt{2\pi}}\int_{h'}^\infty e^{-\frac{1}{2}x'^2}(x'-h')^p\,dx', \text{ if } x' = x/\sigma,\ h' = h/\sigma,$$

$$= N\sigma^p\left(\mu_p(h') - ph'\mu_{p-1}(h') + \frac{p(p-1)}{1.2}h'^2\mu_{p-2}(h') - \text{etc.}\right),$$

where $\mu_p(h')$ is the incomplete normal moment function of the $p$th order. Representing it by $\mu_p'$ for brevity, we have

$$n = N\mu_0', \quad n\nu_1' = N\sigma(\mu_1' - h'\mu_0'),$$

$$n\nu_2' = N\sigma^2(\mu_2' - 2h'\mu_1' + h'^2\mu_0').$$

Hence we find

$$\Sigma^2/\nu_1'^2 = \psi_1 = (\nu_2' - \nu_1'^2)/\nu_1'^2 = (\mu_2'\mu_0' - \mu_1'^2)/(\mu_1' - h'\mu_0')^2 \ldots\ldots\ldots\ldots(1).$$

Now $\psi_1$ is a known quantity, being the ratio of the squared standard deviation of the tail to the squared distance of its mean from the stump. Accordingly if the value of $\psi_1$ be tabled to each $h'$, then,—$\mu_0'$ being known from a table of the probability integral, and $\mu_1'$ and $\mu_2'$ from tables of the incomplete normal moment functions—we shall be able to find $h'$ from the known value of $\psi_1$.

In the next place: $N = n/\mu_0' \ldots\ldots\ldots\ldots(2)$

will be known from the probability integral table of $\mu_0'$ as soon as $h'$ is known.

Lastly: $\sigma = \nu_1'\mu_0'/(\mu_1' - h'\mu_0') = \nu_1'\psi_2 \ldots\ldots\ldots\ldots(3).$

Or, if $\psi_2 = \mu_0'/(\mu_1' - h'\mu_0')$ be tabled for values of $h'$, $\sigma$ will be given at once in terms of $\nu_1'$. Then $h = h'\sigma$ is determined, and the constants of the complete normal distribution found from the constants for the truncated tail.

We require accordingly tables of $\psi_1$ and $\psi_2$. It was found impossible to calculate these with sufficient accuracy from the usual 7-figure tables of the probability integral and the new tables of the normal moment functions. When $h'$ gets at all considerable $\psi_1$ and $\psi_2$ depend on the differences of very small quantities. Accordingly $\mu_1'$ and $\mu_2'$ were calculated *de novo*, and owing to the kindness of Mr Sheppard an unpublished 9-figure table of the probability integral calculated by him was used. Even thus, it has not been considered desirable to give the table to more than three decimal figures.

TABLE I. *Values of the Incomplete Normal Moment Function* $\mu_n(x)$.

A. *Odd Moments* $m_n(x) = \mu_n(x)/\{(n-1)(n-3)(n-5)\ldots 2\}$.

| $x$ | $m_1(x)$ | $m_3(x)$ | $m_5(x)$ | $m_7(x)$ | $m_9(x)$ |
|---|---|---|---|---|---|
| 0·0 | ·0000000 | ·0000000 | ·0000000 | ·0000000 | ·0000000 |
| 0·1 | ·0019897 | ·0000050 | ·0000000 | ·0000000 | ·0000000 |
| 0·2 | ·0078996 | ·0000787 | ·0000005 | ·0000000 | ·0000000 |
| 0·3 | ·0175545 | ·0003920 | ·0000059 | ·0000001 | ·0000000 |
| 0·4 | ·0306721 | ·0012105 | ·0000321 | ·0000006 | ·0000000 |
| 0·5 | ·0468770 | ·0028688 | ·0001183 | ·0000037 | ·0000001 |
| 0·6 | ·0657177 | ·0057372 | ·0003390 | ·0000151 | ·0000005 |
| 0·7 | ·0866883 | ·0101861 | ·0008146 | ·0000493 | ·0000024 |
| 0·8 | ·1092507 | ·0165494 | ·0017172 | ·0001350 | ·0000086 |
| 0·9 | ·1328570 | ·0250925 | ·0032702 | ·0003242 | ·0000259 |
| 1·0 | ·1569716 | ·0359862 | ·0057399 | ·0006988 | ·0000687 |
| 1·1 | ·1810901 | ·0492895 | ·0094199 | ·0013795 | ·0001634 |
| 1·2 | ·2047562 | ·0649423 | ·0146092 | ·0025293 | ·0003549 |
| 1·3 | ·2275737 | ·0827672 | ·0215865 | ·0043539 | ·0007135 |
| 1·4 | ·2492148 | ·1024819 | ·0305828 | ·0070957 | ·0013414 |
| 1·5 | ·2694247 | ·1237174 | ·0417570 | ·0110219 | ·0023776 |
| 1·6 | ·2880214 | ·1460428 | ·0551764 | ·0164068 | ·0040005 |
| 1·7 | ·3048932 | ·1689923 | ·0708039 | ·0235098 | ·0064248 |
| 1·8 | ·3199921 | ·1920929 | ·0884945 | ·0325513 | ·0098944 |
| 1·9 | ·3333265 | ·2148899 | ·1080009 | ·0436894 | ·0146688 |
| 2·0 | ·3449513 | ·2369694 | ·1289874 | ·0569995 | ·0210055 |
| 2·1 | ·3549587 | ·2579749 | ·1510502 | ·0724606 | ·0291380 |
| 2·2 | ·3634677 | ·2776192 | ·1737425 | ·0899486 | ·0392533 |
| 2·3 | ·3706152 | ·2956902 | ·1966019 | ·1092390 | ·0514703 |
| 2·4 | ·3765478 | ·3120515 | ·2191769 | ·1300173 | ·0658224 |
| 2·5 | ·3814140 | ·3266380 | ·2410506 | ·1518971 | ·0822459 |
| 2·6 | ·3853593 | ·3394489 | ·2618602 | ·1744437 | ·1005767 |
| 2·7 | ·3885213 | ·3505370 | ·2813106 | ·1972006 | ·1205553 |
| 2·8 | ·3910268 | ·3599983 | ·2991823 | ·2197160 | ·1418391 |
| 2·9 | ·3929897 | ·3679593 | ·3153329 | ·2415682 | ·1640231 |
| 3·0 | ·3945104 | ·3745671 | ·3296946 | ·2623860 | ·1866637 |
| 3·1 | ·3956755 | ·3799784 | ·3422662 | ·2818638 | ·2093055 |
| 3·2 | ·3965582 | ·3843517 | ·3531029 | ·2997718 | ·2315079 |
| 3·3 | ·3972197 | ·3878403 | ·3623049 | ·3159582 | ·2528687 |
| 3·4 | ·3977101 | ·3905878 | ·3700046 | ·3303476 | ·2730432 |
| 3·5 | ·3980696 | ·3927244 | ·3763548 | ·3429335 | ·2917571 |
| 3·6 | ·3983304 | ·3943653 | ·3815183 | ·3537687 | ·3088145 |
| 3·7 | ·3985175 | ·3956099 | ·3856585 | ·3629529 | ·3240979 |
| 3·8 | ·3986503 | ·3965425 | ·3889331 | ·3706199 | ·3375646 |
| 3·9 | ·3987436 | ·3972329 | ·3914881 | ·3769253 | ·3492376 |
| 4·0 | ·3988085 | ·3977378 | ·3934552 | ·3820351 | ·3591947 |
| 4·1 | ·3988530 | ·3981028 | ·3949499 | ·3861165 | ·3675554 |
| 4·2 | ·3988833 | ·3983635 | ·3960708 | ·3893304 | ·3744677 |
| 4·3 | ·3989037 | ·3985475 | ·3969007 | ·3918258 | ·3800964 |
| 4·4 | ·3989173 | ·3986759 | ·3975073 | ·3937367 | ·3846117 |
| 4·5 | ·3989263 | ·3987645 | ·3979452 | ·3951801 | ·3881809 |
| 4·6 | ·3989321 | ·3988248 | ·3982573 | ·3962557 | ·3909614 |
| 4·7 | ·3989359 | ·3988656 | ·3984770 | ·3970466 | ·3930967 |
| 4·8 | ·3989383 | ·3988927 | ·3986298 | ·3976205 | ·3947135 |
| 4·9 | ·3989398 | ·3989106 | ·3987348 | ·3980315 | ·3959207 |
| 5·0 | ·3989408 | ·3989222 | ·3988061 | ·3983221 | ·3968097 |
| ∞ | ·3989423 | ·3989423 | ·3989423 | ·3989423 | ·3989423 |

TABLE I.—(*continued*).

B. *Even Moments* $m_n(x) = \mu_n(x)/[(n-1)(n-3)(n-5)\ldots 1]$.

| $x$ | $m_2(x)$ | $m_4(x)$ | $m_6(x)$ | $m_8(x)$ | $m_{10}(x)$ |
|---|---|---|---|---|---|
| 0·0 | ·0000000 | ·0000000 | ·0000000 | ·0000000 | ·0000000 |
| 0·1 | ·0001325 | ·0000002 | ·0000000 | ·0000000 | ·0000000 |
| 0·2 | ·0010512 | ·0000084 | ·0000000 | ·0000000 | ·0000000 |
| 0·3 | ·0034951 | ·0000626 | ·0000008 | ·0000000 | ·0000000 |
| 0·4 | ·0081136 | ·0002572 | ·0000058 | ·0000001 | ·0000000 |
| 0·5 | ·0154298 | ·0007604 | ·0000270 | ·0000008 | ·0000001 |
| 0·6 | ·0258121 | ·0018200 | ·0000925 | ·0000037 | ·0000001 |
| 0·7 | ·0394585 | ·0037575 | ·0002588 | ·0000139 | ·0000006 |
| 0·8 | ·0563914 | ·0069507 | ·0006223 | ·0000437 | ·0000025 |
| 0·9 | ·0764632 | ·0118045 | ·0013297 | ·0001177 | ·0000086 |
| 1·0 | ·0993740 | ·0187171 | ·0025857 | ·0002812 | ·0000251 |
| 1·1 | ·1246965 | ·0280428 | ·0046525 | ·0006094 | ·0000658 |
| 1·2 | ·1519070 | ·0400559 | ·0078427 | ·0012160 | ·0001558 |
| 1·3 | ·1804203 | ·0549214 | ·0125028 | ·0022617 | ·0003386 |
| 1·4 | ·2096248 | ·0726741 | ·0189894 | ·0039577 | ·0006842 |
| 1·5 | ·2389164 | ·0932091 | ·0276408 | ·0065653 | ·0012964 |
| 1·6 | ·2677274 | ·1162835 | ·0387442 | ·0103869 | ·0023209 |
| 1·7 | ·2955511 | ·1415300 | ·0525059 | ·0157516 | ·0039494 |
| 1·8 | ·3219594 | ·1684803 | ·0690258 | ·0229926 | ·0064207 |
| 1·9 | ·3466134 | ·1965937 | ·0882796 | ·0324204 | ·0100147 |
| 2·0 | ·3692680 | ·2252921 | ·1101113 | ·0442938 | ·0150415 |
| 2·1 | ·3897700 | ·2539927 | ·1342371 | ·0587910 | ·0218224 |
| 2·2 | ·4080525 | ·2821413 | ·1602593 | ·0759866 | ·0306667 |
| 2·3 | ·4241237 | ·3092387 | ·1876903 | ·0958345 | ·0418437 |
| 2·4 | ·4380556 | ·3348616 | ·2159821 | ·1181613 | ·0555560 |
| 2·5 | ·4499695 | ·3586763 | ·2445598 | ·1426700 | ·0719132 |
| 2·6 | ·4600231 | ·3804450 | ·2728554 | ·1689546 | ·0909136 |
| 2·7 | ·4683965 | ·4000247 | ·3003387 | ·1965228 | ·1124320 |
| 2·8 | ·4752816 | ·4173616 | ·3265431 | ·2248263 | ·1362197 |
| 2·9 | ·4808719 | ·4324798 | ·3510842 | ·2532933 | ·1619132 |
| 3·0 | ·4853546 | ·4454679 | ·3736720 | ·2813629 | ·1890538 |
| 3·1 | ·4889053 | ·4564647 | ·3941138 | ·3085150 | ·2171145 |
| 3·2 | ·4916838 | ·4656432 | ·4123121 | ·3342962 | ·2455315 |
| 3·3 | ·4938321 | ·4731975 | ·4282552 | ·3583379 | ·2737379 |
| 3·4 | ·4954736 | ·4793298 | ·4420056 | ·3803672 | ·3011962 |
| 3·5 | ·4967130 | ·4842409 | ·4536843 | ·4002102 | ·3274261 |
| 3·6 | ·4976381 | ·4881218 | ·4634555 | ·4177877 | ·3520261 |
| 3·7 | ·4983205 | ·4911484 | ·4715111 | ·4331061 | ·3746880 |
| 3·8 | ·4988183 | ·4934784 | ·4780568 | ·4462441 | ·3952025 |
| 3·9 | ·4991771 | ·4952491 | ·4833001 | ·4573366 | ·4134583 |
| 4·0 | ·4994330 | ·4965779 | ·4874418 | ·4665592 | ·4294345 |
| 4·1 | ·4996133 | ·4975627 | ·4906683 | ·4741120 | ·4431886 |
| 4·2 | ·4997391 | ·4982835 | ·4931479 | ·4802063 | ·4548407 |
| 4·3 | ·4998258 | ·4988045 | ·4950279 | ·4850521 | ·4645574 |
| 4·4 | ·4998849 | ·4991766 | ·4964343 | ·4888500 | ·4725352 |
| 4·5 | ·4999247 | ·4994392 | ·4974729 | ·4917846 | ·4789861 |
| 4·6 | ·4999512 | ·4996222 | ·4982298 | ·4940207 | ·4841246 |
| 4·7 | ·4999688 | ·4997483 | ·4987744 | ·4957010 | ·4881574 |
| 4·8 | ·4999802 | ·4998342 | ·4991613 | ·4969464 | ·4912765 |
| 4·9 | ·4999876 | ·4998919 | ·4994326 | ·4978572 | ·4936544 |
| 5·0 | ·4999923 | ·4999303 | ·4996206 | ·4985144 | ·4954417 |
| ∞ | ·5000000 | ·5000000 | ·5000000 | ·5000000 | ·5000000 |

TABLE II.

*Values of the Functions $\psi_1$ and $\psi_2$ required to determine the Constants of a Normal Frequency Distribution from the Moments of its Truncated Tail.*

| $h'$ | $\psi_1$ | $\psi_2$ | $h'$ | $\psi_1$ | $\psi_2$ |
|---|---|---|---|---|---|
| 0·1 | ·588 | 1·311 | 1·6 | ·787 | 2·358 |
| 0·2 | ·605 | 1·371 | 1·7 | ·796 | 2·437 |
| 0·3 | ·622 | 1·432 | 1·8 | ·804 | 2·517 |
| 0·4 | ·638 | 1·495 | 1·9 | ·813 | 2·598 |
| 0·5 | ·653 | 1·560 | 2·0 | ·820 | 2·679 |
| 0·6 | ·668 | 1·626 | 2·1 | ·828 | 2·762 |
| 0·7 | ·682 | 1·693 | 2·2 | ·835 | 2·845 |
| 0·8 | ·696 | 1·762 | 2·3 | ·842 | 2·929 |
| 0·9 | ·709 | 1·833 | 2·4 | ·848 | 3·013 |
| 1·0 | ·722 | 1·904 | 2·5 | ·854 | 3·098 |
| 1·1 | ·734 | 1·977 | 2·6 | ·860 | 3·184 |
| 1·2 | ·746 | 2·051 | 2·7 | ·866 | 3·270 |
| 1·3 | ·757 | 2·126 | 2·8 | ·871 | 3·357 |
| 1·4 | ·767 | 2·202 | 2·9 | ·876 | 3·445 |
| 1·5 | ·777 | 2·280 | 3·0 | ·880 | 3·532 |

This is quite sufficient for practical purposes, for the difficulties of practice do not arise from the paucity of figures in this table, but from quite other considerations, now to be described. In the first place, when we have a truncated distribution like that indicated in the above figure, the group of maximum frequency is at one end of the distribution, and any variation in this frequency widely modifies not only $\Sigma$ but $\nu_1'$. Thus the probable error of random sampling in the "stump" group is very influential. Next, the frequency is almost certain to be given in definite ranges, and the correction of the raw moments for such a case with the maximum frequency group at the terminal is tedious even if it can yet be considered satisfactorily determined. Thus in actual statistics, I have found that truncating at different frequency groups much modifies the values of the unselected frequency constants. For these practical reasons no very great number of decimal places is needful in $\psi_1$ and $\psi_2$. I think, however, the tables may be of service in determining the constants of the untruncated distribution to a first approximation, especially if the final values of those constants be based on truncating at two or three points and averaging the results. Further they may serve to give the *first* idea of the root of the nonic required when we break up a distribution into two components, the constants of one component being indicated by the tail of the frequency on that side.

I have to thank Dr A. Lee for the laborious arithmetic involved in the preparation of the tables accompanying this paper.

A further use of the incomplete normal moment tables in evaluating the incomplete $B$- and $\Gamma$-functions will be considered in another paper.

# ON INHERITANCE OF THE DEFORMITY KNOWN AS SPLIT-FOOT OR LOBSTER-CLAW.

By KARL PEARSON, F.R.S.

(1) My attention was drawn at the beginning of 1907 to the existence of a family in which "split-foot" or "lobster-claw" was an hereditary deformity. I was unaware at the time that Dr Thomas Lewis and Mr Dennis Embleton were at work on a much more elaborate study of the same subject based on quite different material, and when I did know I did not hand over my material to them, as I ought properly to have done, because I wanted to ascertain independently to what extent simple Mendelism really applied to an obviously inherited and fairly simple human deformity. I wanted to convince myself that Mendelism does or does not apply to such cases, by handling the material myself and investigating by all the means at my disposal the authenticity of the records. At the same time my readers will suffer from having nothing like the completeness of detail in my case which may be found in the earlier paper of this number.

The members of the family are scattered through an agricultural district some distance from London, and I could only afford the cost of bringing three members up to London for radiography. In this matter I must mention my deep indebtedness to Miss C. O. Stevens not only for much aid in following up the family history from the clue she first reported to me, but also for the arduous task of piloting the strangers through the sights of the Metropolis. To Dr Mackenzie Davidson I owe the further big debt that, in the interests of science, he gave up a large portion of his valuable time to preparing by stereoscopic radiography no fewer than 22 negatives of the feet and hands of these three individuals.

Only a few of these radiograms are reproduced in this paper, but the variety of positions taken, as well as the stereoscopic nature of the pictures, has enabled my colleague, Professor Thane, to give a very complete account of the hands and feet of these three individuals. The ever ready aid given by Professor Thane to biometric inquiry would be very inadequately expressed by a few lines of thanks in a single biometric memoir.

It would indeed be impossible from Professor Thane's account of these three cases to predict the precise nature of the deformity—the presence or absence of bones—in other members of the family who have not been radiographed. No

great faith is to be put in the family record of the number of perfect fingers or toes possessed by dead or living members. The list given on pp. 74, 75 must only be trusted for such obvious and gross variations in the nature of the deformity as supernumerary digits or syndactylised fingers. The importance of the accurate description of each individual member, however desirable, is less, after the work of Messrs Lewis and Embleton; their memoir fairly indicates the range of types to be expected. These types are largely represented in the members of my family who have been more fully examined; there are certain additional types also. It is clear that what they and I are dealing with is an hereditary deformity of the hands and feet, with a definite, if remarkably wide, field of variation.

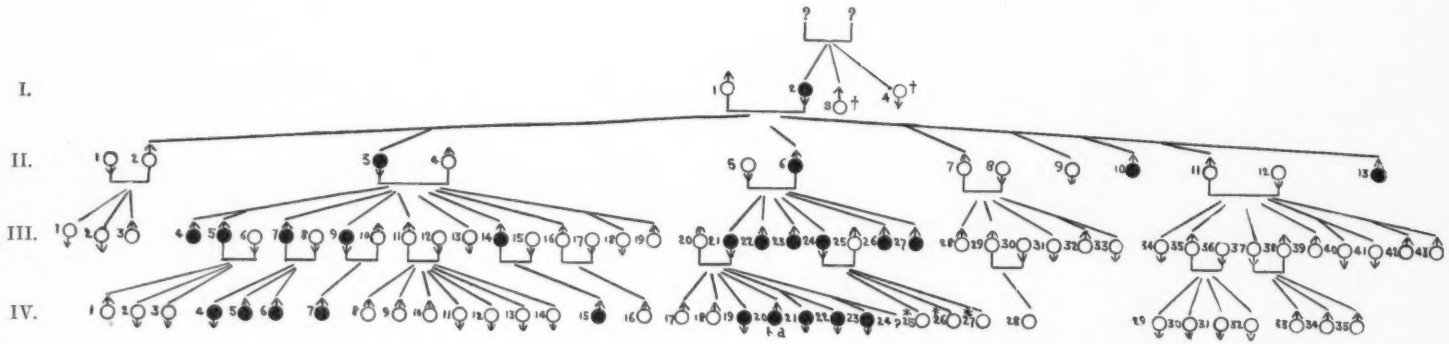
Lastly, I have most heartily to thank Mr J. H. Astbury, whose intimate knowledge of the district and its inhabitants has been most helpful to me in my inquiries.

(2) Before turning to more definite details as to the present family, I should like to mention the existence of a similar family, with names Bell and Agnew, at Whithorn, Wigtownshire, some 50 years ago. An old man told my informant that a boy with deformed hands was a descendant of the "Cleppie Bells." This was a family, one of which had assisted as Sheriff's officer at the drowning of the Wigton Bay martyrs in 1685. On that occasion the officer Bell said to a young maid, Margaret Wilson, "Will you not say: God bless King Charlie, and get this rope from off your neck?" "God bless King Charlie, if He will," she responded. Whereupon he said "Clep down among the partens and be drowned." Thus he was called "Cleppie Bell," and his descendants have ever afterwards suffered from a deformity of the hand, although sometimes a generation is missed over. It is true that the term "clepped" in Scotland refers to webbed, but the reference to the lobsters suggests "lobster-claw" deformity, and one phase of this, the syndactyly, might easily be described as fingers grown or webbed together. The boy referred to above became a sailor and was last seen in Glasgow 20 years ago. No stress whatever can be laid on such a tale as this, but it is interesting as pointing to the existence in Scotland* of a deformed hand inheritance for nearly two hundred years, and the information might be followed up by any one coming across Scottish cases of either webbed hands or lobster-claw. This is the only case I have met with in which there is any statement that deformed offspring were born from undeformed parents†. Throughout both my family and that of

* Messrs Lewis and Embleton state that their family originated in Scotland.

† This follows also from the somewhat different version given in Sir Andrew Agnew's *The Hereditary Sheriffs of Glasgow*, Edinburgh, 1893. We read in Vol. II. p. 142: "Still more grotesque is the tradition of the 'Cleppie Bells.' A constable who was held to have carried out his orders unfeelingly, as he fastened the women to the stakes, was asked how the poor creatures behaved when the cold wave roared and foamed about their heads. 'Oo,' he replied jocularly, 'they just clepped roun' the stobs like partens, and prayed.'—Soon after Bell's wife was brought to bed, when the howdie exclaimed in horror: 'The bairn is clepped!' (i.e. the fingers grew firmly together). Another child was born, and yet another, and as each little wretch in turn was seen to be 'clepped' the most incredulous were convinced it was a judgment of Providence.—We have been gravely assured that within the memory of man a female descendant of the bad constable on giving birth to a child, was horrified by the exclamation, 'The bairn is clepped.'"

FAMILY TREE. INHERITED "LOBSTER CLAW."




† No record of condition of I. 3 and I. 4.

* Died at birth, no record.

Messrs Lewis and Embleton undeformed members give rise only to undeformed offspring. It is not possible to say whether the deformity really is latent or not, there have been no cousin marriages to form even a rough test of latency.

The absence of cousin marriages of any kind compels us to consider "lobster-claw" as a dominant character, and after the ancestress I. 2 of my pedigree, who may in Mendelian terminology possibly have been a homozygote, all the descendants must be looked upon as heterozygotes. On this assumption all normal individuals are recessive, and the fact that normal individuals from tainted stocks do not have deformed offspring meets with a ready explanation.

It is this noteworthy point not only in the present family, but in Lewis and Embleton's family, in Drinkwater's Brachydactylous Family and in Nettleship's Night Blind Family, which is the main-stay in these cases of the Mendelian theory. It is perfectly true that it would flow from other hypotheses, for example from the assumption that the gamete was not pure, but that dominance flowed from a numerical preponderance of allogenic determinants*. Such a theory could only be tested by inbreeding. For the sake of science, therefore, if not for local well-being, it is desirable that marriages between the normal members of these stocks should occur, and this even in sufficient numbers to test whether there are really latent tendencies to the deformity in normal members of the stock. Should this not prove to be the case, the non-marriage of the deformed members would be sufficient to check the spread of the deformity. Obviously a determinantal theory not based upon the pure gamete would further account for the occasional and rare appearance of the deformity in a child of normal parents. Mendelism must assert in a case like the present where the character must be dominant, that such an appearance is a mutation or sport.

(3) Since the deformity must be looked upon as dominant in the Mendelian sense, it follows that on the average half the offspring of a deformed parent ought to be deformed. We have data for such families in three generations, and the following results flow from them:

| N = Normal<br>D = Deformed | Families | | | | | Totals |
|---|---|---|---|---|---|---|
| | *1* | *2* | *3* | *4* | *5* | |
| | N. D. | N. D. | N. D. | N. D. | N. D. | N. D. |
| II. Generation | 4 4 | — | — | — | — | 4 4 |
| III. Generation | 5 5 | 0 6 | — | — | — | 5 11 |
| IV. Generation | 0 3 | 0 1 | 0 1 | 2 5 | 3 0 | 5 10 |
| Totals | — | — | — | — | — | 14 25 |

* See the paper in this number of *Biometrika* on p. 80.

The amount of material here is not very great, but, as far as it goes, it does not support the conclusion that for the tainted families the abnormal members are relative to the normal fewer in number. It seems rather to indicate that in the last generations the abnormal have been twice as numerous as the normal. My data are, however, far less numerous than Messrs Lewis and Embleton's. Leaving out their two doubtful cases, I find (*Biometrika*, Vol. VI. Plate I) that they have 22 normal to 43 abnormal members in the families of their stock with deformed parents. I shall confine my attention to their stock and my stock, because I feel in my own case the confidence of the personal collector, and am sure that a number of individuals have not been omitted or wrongly classed; and because I realise in their case also that Messrs Lewis and Embleton have examined their stock with the special view of testing definite ratios. The families of other observers may have been carefully worked out, but the importance of complete enumeration has only been recognised since it has become desirable to test Mendelian theory.

Examining the data from the statistical standpoint, I draw attention in the first place to the first family of the third generation. This family should approximate to 3 normal to 3 deformed, but every member born was deformed. The odds against a run of six deformed are 63 to 1, if the chances of normality and of lobster-claw are equal. The odds against a deviation as great as 25 to 14 from a ratio of equality are 14 to 1; the odds against Messrs Lewis and Embleton's ratio of 43 to 32 are about 9 to 1; the odds against such an excess as is represented by our combined results 68 to 46,—a ratio corresponding to 3 to 2 instead of 1 to 1,—are more than 49 to 1. Now this excess of the abnormal in any deformity like the present ought to be admitted, if it actually exists, for it is not only opposed to simple Mendelism, but its bearing on Eugenics is of the greatest importance. While I have found the like excess in other cases, it does not appear to be true for all deformities. Thus, from the pedigree plate in Mr Nettleship's admirable paper on Night Blindness*, I deduce the following results:

| | Abnormal | Normal |
|---|---|---|
| Generation III. | 10 | 6 |
| Generation IV. | 15 | 43 |
| Generation V. | 26 | 29 |
| Generation VI. | 30 | 51 |
| Generation VII. | 27 | 55 |
| Generation VIII. | 14 | 36 |
| Generation IX. | 9 | 11 |
| Totals | 131 | 231 |

* *Ophthalmological Society's Transactions*, Vol. XXVII. pp. 269–291.

The odds against such a deviation from equality are enormous, i.e. about 10,000,000 to 1. But it will be noted that the abnormal members are largely in *defect.* Mr Nettleship himself remarks that this deficiency may probably be explained by unavoidable imperfections in the record, especially in the earlier generations. The excess of normals, however, has been more than maintained in the recent generations; and if, in defending the Mendelian ratio of equality, we impugn the record on the ground that members of the stock are ashamed to own their defect and screen it, what then becomes of the evidence it undoubtedly provides for Mendelian theory in the "invariable continuity of descent of the disease in the affected branches, and its permanent disappearance from all other divisions of the genealogy"? The screening of the disease must clearly have become an hereditary character, and the above evidence of segregation would be worthless. This principle of segregation is so vital, is such a wonderful, and I am inclined to say, glorious addition to our knowledge of heredity, that I am inclined in these cases of family abnormalities to defend Mendel against Mendel, or to preserve Mendelian segregation at the expense of Mendelian ratios. We have seen that the odds against a ratio of 3 to 2, which is the combined result of our split-foot families, are very slight, and this applies equally to the case of Night Blindness, where a random sample deviating as much as the present one from the ratio of 3 to 2 would occur once on the average in 6 or 7 trials. There is an additional point also to be noted. The Curé of Vendémean, according to Mr Nettleship, states that if the original night-blind individual, "the Jean Nougaret of 1637 could reappear to-day, except for a few individuals recently established in the district, all could salute him as their ancestor" (*loc. cit.* p. 288). If this be true, then the Mendelian ratio to be expected is more than 1 to 1, for the husband or wife of the night-blind individual would, in certain cases, be screening their own defect, and the offspring ought to have been night-blind in the ratio of 3 to 1. On the whole, while these cases give very definite evidence of the segregation factor, they do not seem to me to favour the segregation in rigid Mendelian proportions.

(4) When we turn to the actual nature of the deformity in the case I am dealing with, we find all the principal types observed by Messrs Lewis and Embleton recur, with one exception which may exist, but owing to the few individuals who have been radiographed, this cannot be asserted. I refer to the existence of cross bones (*Biometrika*, Vol. VI. p. 41). On the other hand, individuals exist showing syndactyly and polydactyly; gross deformity, several "fingers" placed promiscuously on a large misshapen "hand"; or branching fingers which, without careful radiographic examination, it would perhaps be hard to classify as polydactyle or misplaced.

The usual type of hand is to outward appearance, the hook-like form of my Plate XI, but the degree of defectiveness in the bones of the hand, even as affecting the carpus, may vary widely. In one case (II. 3) the hands were perfect but the feet deformed. There seems to be some evidence that in the younger generations the defectiveness is increasing; thus all the children of II. 6 had two-

toed feet, but the children of III. 21 are one-toed, with one exception, who has one two-toed foot, and the other one-toed. In this family as distinct from that of Lewis and Embleton the defectiveness in the foot may reach the tarsus. Contractions of the fingers and toes, these being flexed towards the centre of the member, are common. The following is a rough external account of the deformity of the individuals in the pedigree:

*List of deformed members of the Stock.*

I. 2. Ann J. Nothing known of her parents; said to have had a brother and sister but no record of their condition to be found. She had one finger only to each hand, but the 1st and 5th toes on both feet. Died at 77. There was no consanguinity between I. 1 and I. 2.

II. 6. Two fingers on right hand, one finger and thumb on left; 1st and 5th toes on each foot. No consanguinity with wife II. 5. Cause of death, asthma.

II. 10, II. 13. Deformed, both hands and feet.

II. 3. Hands perfect, both feet only 1st and 5th toes. Husband II. 4 not related.

III. 21—24, 26, 27 are children of II. 6 above.

III. 21. One finger to each hand, and 1st and 5th toes. Subject to fits. Unrelated to husband, III. 20. See Plates XI, XII, XIII and XV, Fig. vii.

III. 22. Right hand six fingers, the supernumerary beyond the little finger; left hand five fingers, all deformed and bent at angles to palm of hand. 1st and 5th toes on each foot. Dead.

III. 23. Both hands one finger, both feet 1st and 5th toes. Diseased arm bone.

III. 24. Right hand ring and little fingers, but these united together; left hand two bent fingers and thumb. 1st and 5th toes on each foot.

III. 25. Right hand two fingers only; left hand two fingers and thumb; both feet 1st and 5th toes.

III. 26. One finger on each hand; 1st and 5th toes on each foot.

III. 4, 5, 7, 9 and 14 are children of II. 3. No details except that both hands and feet are affected.

In the fourth generation, we have first children of III. 21.

IV. 19. Both hands and feet affected; 5th toe only. Married, no offspring at present.

IV. 20. Both hands and feet affected; 5th toe only; stillborn.

IV. 21. Only 5th toe on each foot; little finger on each hand. Has had one or more fits. Plates XI, XIV and XV, Fig. viii.

IV. 22. Both hands and feet affected; 5th toe only; weakly constitution.

IV. 23. One finger on each hand. 1st and 5th toes on right foot, 5th toe only on left foot. See Plates XI and XVI.

IV. 24? Expected Birth*.

IV. 18 has also a weakly constitution though hands and feet are normal.

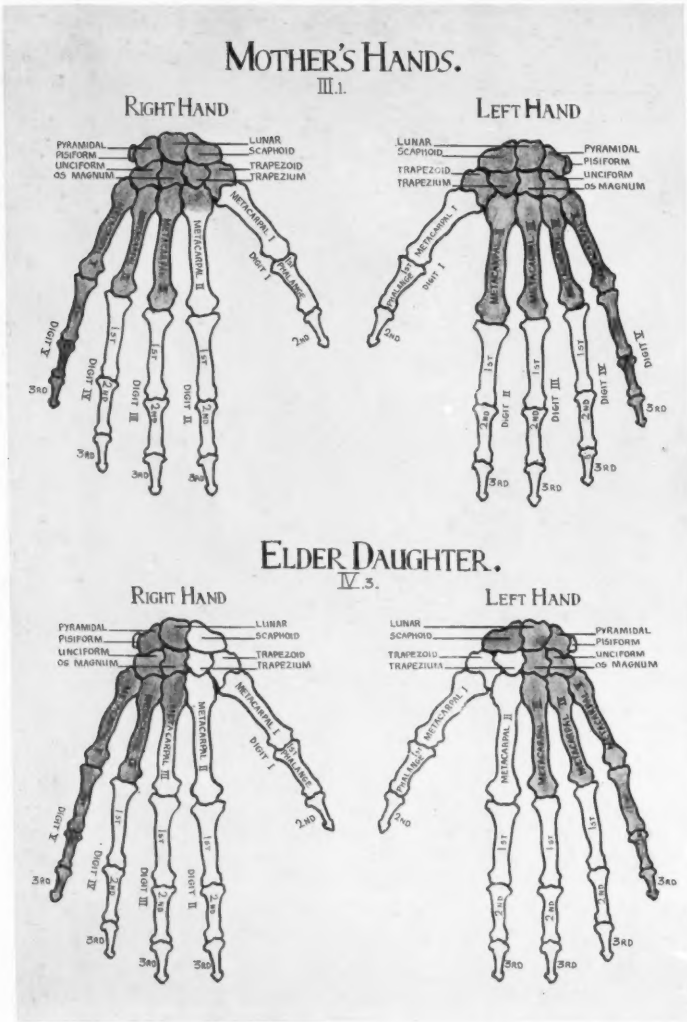
Of the descendants of II. 3, the grandchildren through III. 7 and III. 8 (who are unrelated) are:

IV. 4. Ring and little finger only, joined together on each hand. 1st and 5th toes on each foot.

IV. 5. One little finger only on each hand; 5th toe only on each foot.

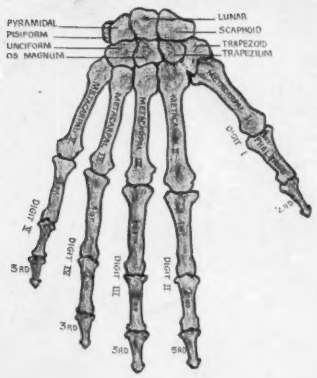
IV. 6. One little finger only on right hand; thumb and little finger on left hand; 1st and 5th toes on each foot.

* Born as this proof goes finally to press; a girl deformed in both hands and feet. Thus the improbability of the Mendelian quarter is greater than that given above.

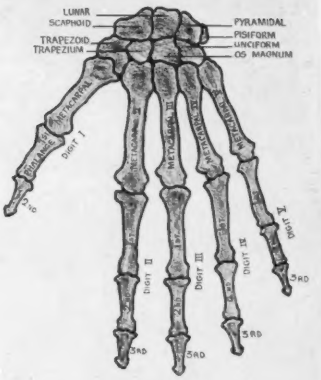

MOTHER'S HANDS.
III.1.
RIGHT HAND
LEFT HAND
PYRAMIDAL
PISIFORM
UNCIFORM
OS MAGNUM
LUNAR
SCAPHOID
TRAPEZOID
TRAPEZIUM
METACARPAL I
METACARPAL II
METACARPAL III
METACARPAL IV
DIGIT I
DIGIT II
DIGIT III
DIGIT IV
PHALANGE
1ST
2ND
3RD
ELDER DAUGHTER.
IV.3.
RIGHT HAND
LEFT HAND

# Normal Hands.

## Right Hand



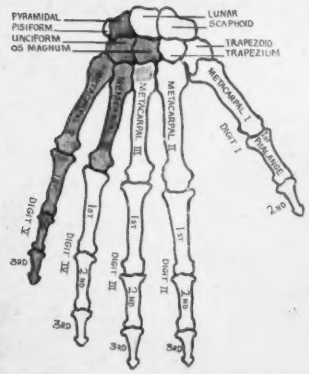

## Left Hand



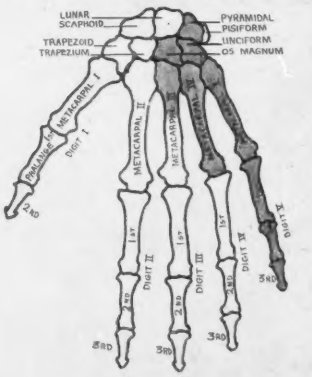

# Younger Daughter.

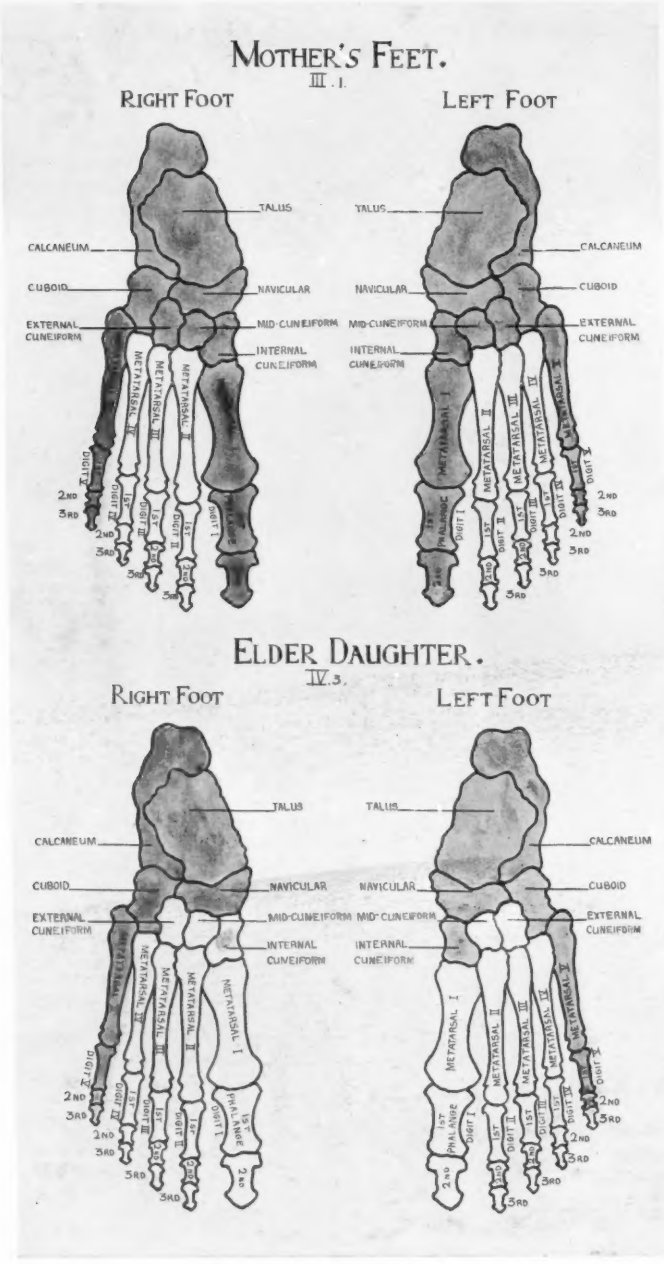
IV.7.

## Right Hand



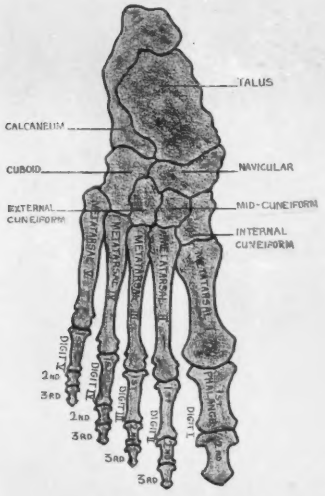

## Left Hand



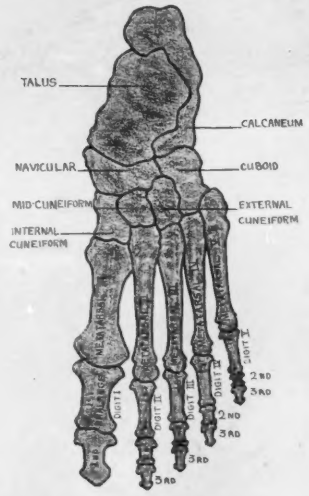

MOTHER'S FEET.
III. 1.
RIGHT FOOT
LEFT FOOT
TALUS
CALCANEUM
CUBOID
NAVICULAR
EXTERNAL CUNEIFORM
MID-CUNEIFORM
INTERNAL CUNEIFORM
METATARSAL I
METATARSAL II
METATARSAL III
METATARSAL IV
METATARSAL V
DIGIT I
DIGIT II
DIGIT III
DIGIT IV
DIGIT V
1ST PHALANGE
2ND
3RD
ELDER DAUGHTER.
IV. 3.
RIGHT FOOT
LEFT FOOT
TALUS
CALCANEUM
CUBOID
NAVICULAR
EXTERNAL CUNEIFORM
MID-CUNEIFORM
INTERNAL CUNEIFORM
METATARSAL I
METATARSAL II
METATARSAL III
METATARSAL IV
METATARSAL V
DIGIT I
DIGIT II
DIGIT III
DIGIT IV
DIGIT V
1ST PHALANGE
2ND
3RD

# Normal Feet.

## Right Foot



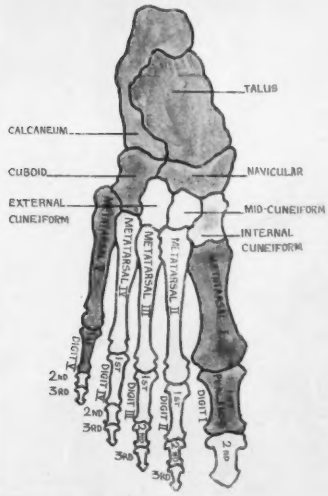

## Left Foot



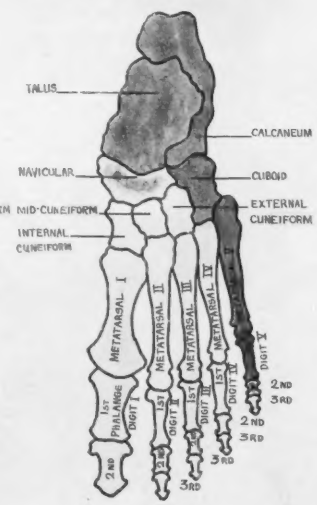

# Younger Daughter.

IV. 7.

## Right Foot




## Left Foot



Of the other two grandchildren of II. 3, belonging to the deformed group, IV. 7 and IV. 15, I only know that both feet and hands are deformed.

Attention must be drawn to the fact that the deformed individual III. 5 has had three children all normal, and to the fact that the deformed individual III. 24 has had three children born dead, of whom nothing is known as to presence or absence of deformity.

(5) Of the three individuals III. 21 and her two daughters IV. 21 and IV. 23, aged respectively 10 and 2½ years, the following account has been drawn up by Professor Thane from Dr Mackenzie Davidson's radiograms. In the list below a plus sign marks the presence of the corresponding bone, the dot its absence. In the diagrams of Plates IX and X the defectiveness of the bones of the hands

| Hands | Mother III. 21 Age 37 | | Elder Daughter IV. 21 Age 10 | | Younger Daughter IV. 23 Age 2½ | | Feet | Mother III. 21 | | Elder Daughter IV. 21 | | Younger Daughter IV. 23 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Right | Left | Right | Left | Right | Left | | Right | Left | Right | Left | Right | Left |
| Scaphoid ... | + | + | • | + | • | • | Calcaneum ... | + | + | + | + | + | + |
| Lunar ... | + | + | + | + | • | • | Talus ... ... | + | + | + | + | + | + |
| Pyramidal ... | + | + | + | + | + | + | Navicular ... | + | + | + | + | + | + |
| Pisiform ... | + | + | • | • | • | • | Intl. Cuneiform | + | + | + | + | + | • |
| Trapezium ... | +? | + | • | • | • | • | Mid. Cuneiform | + | +? | • | • | • | • |
| Trapezoid ... | + | + | • | • | • | • | Extl. Cuneiform | + | + | • | • | • | • |
| Magnum ... | + | + | +? | + | + | + | Cuboid ... | + | + | + | + | + | + |
| Unciform ... | + | + | + | + | + | + | Metatarsal 1 ... | + | + | • | • | + | • |
| Metacarpal 1 | • | • | • | • | • | • | „ 2 ... | • | • | • | • | • | • |
| „ 2 | + | + | • | • | • | • | „ 3 ... | • | • | • | • | • | • |
| „ 3 | + | + | + | + | +? | + | „ 4 ... | • | • | • | • | • | • |
| „ 4 | + | + | + | + | + | + | „ 5 ... | + | + | + | + | + | + |
| „ 5 | + | + | + | + | + | + | Digit 1 ... | + | + | • | • | + | • |
| Digit 1 | • | • | • | • | • | • | | 2 ph. fused | 2 ph. | | | 1 ph. | |
| „ 2 | • | • | • | • | • | • | | | | | | | |
| „ 3 | • | • | • | • | • | • | „ 2 ... | • | • | • | • | • | • |
| „ 4 | • | • | • | • | • | • | „ 3 ... | • | • | • | • | • | • |
| „ 5 | + | + | + | + | + | + | „ 4 ... | • | • | • | • | • | • |
| | 3 ph. | 3 ph. | 3 ph. | 3 ph. | 3 ph. | 3 ph. | „ 5 ... | + | + | + | + | + | + |
| | | | | | | | | 3 ph. | 3 ph. | ?2 ph. | 2 ph | 1 ph. | 2 ph.? |

and feet is indicated by noting the presence of bones by shading in each case, the unshaded bones are absent, a key figure indicates the bones of the normal hand and foot*. These figures show nothing of the contracted and distorted form of the original members, for which the reader is referred to Plates XII to XVI which give a few of the radiograms of these three cases. It is fairly possible by the shadow of the fleshy parts and the aid of Plate XI to reach a reasonable appreciation of the actual appearance of the living individuals.

The nature of the deformity in these limbs is a partial suppression of the digits, together with reduction (absence or fusion) of certain associated carpal or

* Owing to an oversight of the draughtsman, the names of the bones were inserted on *all* the figures; it seemed better to leave the redundancy than to go to the labour of redrawing the plates.

tarsal elements in the more pronounced cases. The bones of the forearm and leg are normal in their disposition as far as they appear in the photographs.

In the mother, III. 21, the left hand has a well-developed carpus which may be regarded as normal, while in the right hand the distal row of the carpus shows only three bones, the outermost of which probably represents the fused trapezoid and trapezium. In both hands the fifth digit (little finger) is well developed; the whole thumb is missing, as are also the phalanges of the index, middle and ring fingers, those digits being represented only by the metacarpal bones, of which the proper second (that of the index finger) is the smallest, and indeed in the right hand is reduced to a small piece of bone corresponding to the basal portion only of the normal bone.

In the left hand of the daughter, IV. 21, there are three bones in the proximal row of the carpus, the pisiform being as yet unossified, two only in the distal row, magnum and unciform, the three inner metacarpal bones, diminishing in size from the fifth to the third, and the normal three phalanges of the little finger. Suppressed are the two outer bones of the distal row of the carpus, trapezium and trapezoid, the first and second metacarpal bones, and the phalanges of the thumb, index, middle and ring fingers. The right hand shows a greater degree of reduction, the scaphoid being suppressed in the proximal row of the carpus, while in the distal row there is only one bone representing the fused magnum and unciform, and the third metacarpal is smaller than in the left hand.

IV. 23 has only three carpal bones, pyramidal, magnum and unciform, in each hand, but she is still too young for ossification to have started in the others; the fourth and fifth metacarpal bones are fairly well developed; the third metacarpal is small, in the right hand merely a nodular vestige of the basal portion; the phalanges of the little finger are present and of good size.

The malformation of the hands consists therefore of suppression or reduction of parts proceeding from the radial (thumb) side towards the ulnar (little finger) side, and carried to the greatest extent in the young child, IV. 23.

The feet are not so homogeneous as the hands, and fall into two well-marked groups. In the one, comprising the two feet of the mother and the right foot of the younger child, IV. 23, the three middle digits are suppressed, including metatarsal bones and phalanges, while the first and fifth are fairly well developed, especially in the mother. The tarsus may be regarded as complete in the mother, only in the right foot the number of elements is apparently reduced by the fusion of the internal and middle cuneiform bones; while in the right foot of IV. 23 the middle and external cuneiform bones are missing, but it may be that they are still in the cartilaginous stage, ossification not having yet begun.

The two feet of IV. 21 and the left foot of IV. 23, which constitute the second group, are also very much alike. Here the inner four digits are suppressed, and only the fifth (little) toe is represented. In all the talus, calcaneum and cuboid of the tarsus are well developed, the navicular is ossified (just beginning to ossify in

IV. 23)*, and the external and middle cuneiform are wanting, while a rudimentary internal cuneiform is present in both feet of IV. 21; that this is not seen in IV. 23 is again possibly due to delayed ossification.

Hence the second group differs from the first in the suppression also of the first digit, and probably a concomitant reduction of the tarsus. The close resemblance of the condition in this second group to that exhibited by the hands is obvious.

All through the deformity is greater in the children than in the mother.

In the hands, IV. 23, the younger daughter, shows greater degree of reduction than IV. 21. The small number of carpal bones in IV. 23 is to be explained however by the fact that the child is too young for ossification to have begun in the missing elements. The same consideration will account for the absence of the pisiform in IV. 21. This does not apply to the metacarpal bones.

In the feet there is a difference on the two sides. The left foot is slightly more reduced in IV. 23 than in IV. 21; whereas the right foot is much more defective in IV. 21, IV. 23 having the first digit fairly well developed.

In all three the right hand shows a greater degree of reduction than the left, expressed in the condition of the metacarpal bones.

In the mother the left foot may be regarded as being slightly more reduced, as indicated by the fusion of the internal and middle cuneiform bones. In IV. 21 the reduction is slightly greater in the right foot than in the left; whereas in IV. 23 the left foot is the most reduced of the whole series, while the right foot has a well developed metatarsal and phalanx to both great and little toes.

(6) I have delayed referring to my final criticism of the application of Mendelian conceptions to cases of abnormality like the present, until I had placed before the reader the above accurate account of the nature of the defect in two or three individuals. But having done so it seems to me that a very difficult question has to be answered. What in a case like this of split hand and foot is the Mendelian unit character? What are the allelomorphs whose combined absence or combined presence is impossible in the zygote? We can understand what is meant when an organ is said to be yellow or white, rough or smooth, however much we may wish for the use of a quantitative scale of such characters. We can understand what is meant by the presence or absence of a given individual bone, say the second phalanx of the fifth digit. But in a case like the present we are concerned not with a simple unit of any kind, but the absence of a complex of even as many as 60 bones situated in four different parts of the body. In other cases 10, 20 or 40 bones may be wanting or incompletely developed. It is true that these parts are highly correlated, but as we have seen, deformity of the feet is not always accompanied by deformity of the hands. Can we look upon the character here as built up of a series of coupled characters representing the system of bones we have referred to? If so, how can we explain the variety of detail which is to be found within the same family? How shall we account for other

* It is interesting to the anatomist to note that the tuberosity of the navicular bone has a distinct centre of ossification in both feet of IV. 21.

families in which mere syndactyly or polydactyly occurs, or again in which club-foot alone is inherited? It appears to me that in a case like the present we cannot possibly speak of the presence or absence of a unit character, there must be an inertia or weakening of some development-controlling determinant. I will not venture to call it an absence of such a determinant, because its range varies so widely from individual to individual. If we use the word absence we must make the full control depend on the presence of a large number of units, and there must be—as for example in the reducing division—some process by which a random number of these are cast out. Even with this interpretation we are really thrown back on quantitative degrees in the controlling factor. These control units cannot be specific, for if they were associated with special bones, it is not easy to understand how a bone occurring in parent and grandparent* could disappear in the zygote resulting from a normal union. The problem has been forced on me recently from the practical side in studying cases of human degeneracy with correlated defects, and although I have thought over it carefully I can see no daylight theoretically in these questions of correlated physical or mental defects —defects having a considerable range of variation, but peculiar to certain stocks —except from the assumption of development-controlling units or determinants.

These determinants may be more or less definitely associated with certain portions of the body—hands and feet, or eyes for example—but the total number rather than any specific character of each must determine in every individual case the extent of defective development. Whatever the *modus* may be, i.e. whether the determinantal system be quantitatively continuous or consist of discrete units, I venture to suggest that it is rather in the direction of inertness in controlling determinants than in the hypothesis of a definite allelomorphic character in the Mendelian sense that we must seek for light on the inheritance of physical or mental degeneracy or deformity in man. The complexity and variety of the deformity in many of these cases is too great for us to attribute it to a single allelomorph, and there is thus no obvious reason for anticipating a simple Mendelian ratio.

(7) On the subjects of my present study a few general remarks may be made. They show remarkable aptitude in the use of their misshapen hands. At school the children hold their own with normal scholars in writing, drawing, and even needlework. The family rather resent anything in the shape of pity. One of its men expressed himself as follows: "Bless 'e, sir, the kids don't mind it. They never had the use o' fingers and toes, and so they never misses 'em." One of the girls with the 5th digit only on each hand is learning dressmaking as a pursuit. In Messrs Lewis and Embleton's family the gait appears to be normal†, but I should hardly say that this describes those members of the present family that I have seen. One who has long been acquainted with them writes:

* E.g. the first digit of the foot existing in III. 21 and II. 2 disappears in the offspring of a marriage of III. 21 with a normal, as in IV. 21.

† *Biometrika*, Vol. VI. p. 28.

Hands of Mother (III. 21) and Daughters (IV. 21 and IV. 23)

Fig. ii

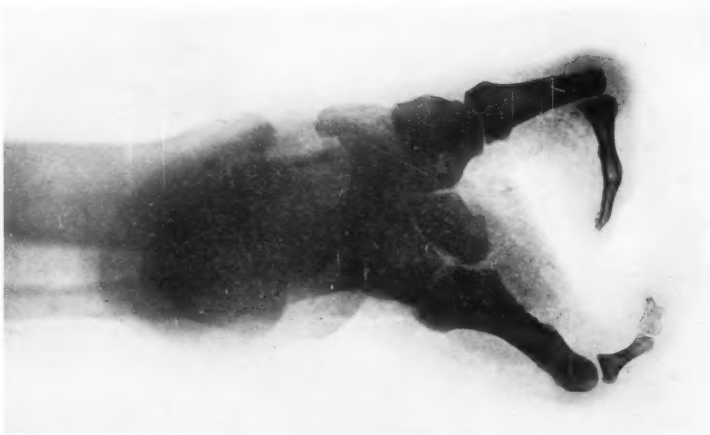


Fig. i

Feet of Mother (III. 21). Front view

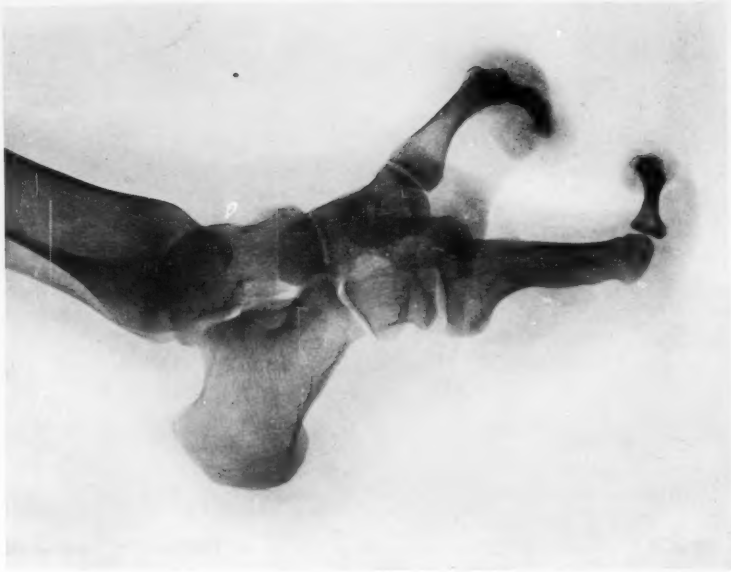


Fig. iv

Fig. iii Feet of Mother (III. 21). Side view

Fig. v

Fig. vi

Feet of elder Daughter (IV. 19). Front and side views

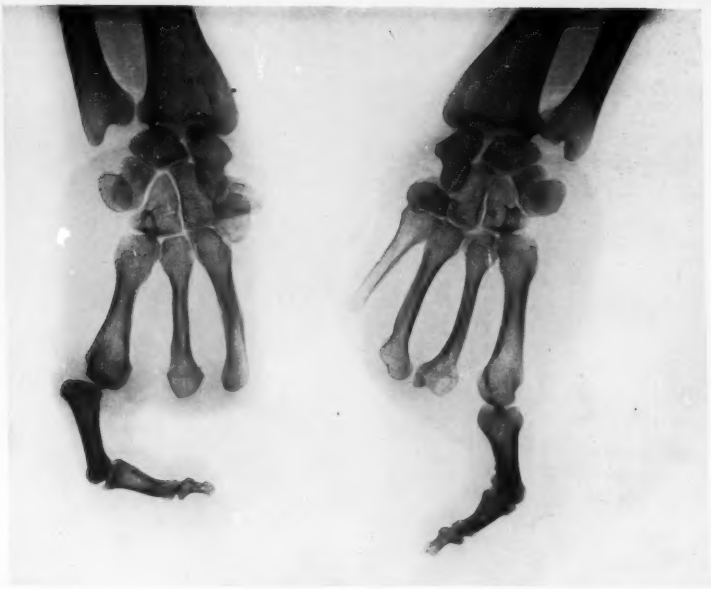


Fig. vii

Hands of Mother (III. 21)

Fig. viii

Hands of elder Daughter (IV. 19)

Fig. ix

Hands of younger Daughter (IV. 23)

Fig. x

Feet of younger Daughter (IV. 23)

"Their feet lack the natural spring of the normal foot and their gait is therefore very ungainly, being assisted by a peculiar swing of the arms, much as a normal person would walk on heels alone. This is of course a great drawback, but not such as to incapacitate them from earning a living."

Turning to a further point we note that the deformed members seem to have no reduced fertility, nor do they appear to have difficulty in finding normal husbands or wives. On this account there is some concern at the perpetuation of the deformity in the district. Physically we may perhaps detect in some members signs of degeneracy, but they cannot be said as a rule to have weak constitutions. Intellectually the children take high rank in intelligence, or perhaps it would be better to say in "cuteness." Thus physically handicapped it would perhaps be unreasonable to expect in all members the highest social qualities, and assuming the stock to have been initially good, it may be doubted whether the deformity permits its members to select the physically and mentally fittest of normal mates. The stock is sufficiently extensive now to intermarry, and should it be ostracised might do so, a result which would have much scientific interest; but which might lead to the perpetuation of a split-foot race. Eugenically the problems associated with the family are not only scientifically difficult but practically serious; we live no longer in an age when the need for perfect hand and foot in the struggle for existence preserved the race from the perpetuation of a deformity. How is protection now to be maintained? In view of possible future legislation, it would seem for national purposes desirable that stocks of the present type should be kept under observation and careful records preserved of the history of all their branches. One of the most urgent social necessities is that every birth certificate should be associated with a medical certificate of the normality or abnormality of the child registered. A separate filing of the abnormal cases would not only keep the scientist in touch with important material, but provide the legislator with data whereby to appreciate the extent to which the suspension of natural selection is racially deleterious.

# ON A MATHEMATICAL THEORY OF DETERMINANTAL INHERITANCE, FROM SUGGESTIONS AND NOTES OF THE LATE W. F. R. WELDON.

By KARL PEARSON.

(1) *Introductory.* In the memoir of Weldon's life published in the last volume of this Journal* occurs the following paragraph referring to the year 1905:

"In the summer the present writer was at East Ilsley, some seventeen miles from Oxford, and there was cycling out several times a week; the writer's chief work was on other than biometric lines and broken by other claims on his time, but there was steady joint work on the determinantal theory of inheritance as outlined by Weldon, and it is hoped that it is sufficiently advanced to be completed and published."

That hope is not wholly fulfilled. My mathematical draft of the theory was finally taken by Weldon from Ilsley to Oxford to be rewritten with proper biological terminology, and to have erroneous developments marked for reworking; this revision never appears to have been accomplished, and beyond my attempt at mathematical interpretation of Weldon's ideas, which I knew to be inapplicable at certain points, and his letters and a notebook full of numerical illustrations due to an earlier period of the same year, I have practically nothing to guide me in reconstructing Weldon's theory. The possibilities before me were threefold, to allow the matter to drop, to wait till I could find a cytologist with an interest in and a knowledge of the theory of chance at all comparable with Weldon's before going further, or lastly to publish what I could put together on my own responsibility. I have adopted the latter course, not because I am unconscious that I am likely to blunder, but because the material may at some time prove sufficiently suggestive to a reader with competent knowledge to cause him to take up the subject on broader and more effective lines, than I can pretend to do.

Accordingly for all that is suggestive or valuable in the present discussion Weldon is responsible. My contribution is solely the mathematical analysis, and even here directly, and by force of numerical illustration, he had gone a considerable way independently. Any defects in terminology and misstatements of cytological facts are mine alone.

* Vol. v. p. 46.

In starting his investigations Weldon proceeded on very general lines, which it may be well to indicate. He was not inclined to accept the theory of unit characters, of allelomorphs and of pure gametes as capable of fully describing even the inheritance of the simplest characteristics. At the same time he recognised the importance of the segregation first pointed out by Mendel in the offspring of hybrids, though even here he was not prepared to make the segregating classes so distinct and so wanting in continuous variation, as some Mendelians have held them to be. He was convinced that in a sufficiently general theory of inheritance some place must be left for a normally arising percentage, however small, of variants, specially related to the distant ancestry. He was thus seeking for a mechanical explanation of latent characters, or in other aspects of reversions and even mutations.

In starting his work Weldon was perfectly catholic as to the possibility or not of paternal and maternal chromosomes retaining their individuality from the moment of fertilisation up to the reducing division of the germ cells. He did not consider this point as at present absolutely settled by the cytologists. While the retention of individuality immensely simplifies the analysis, this did not deter Weldon from investigating at great length the numerical results flowing from supposing an interchange of chromomeres at every mitosis.

The following rough notes of January 2, 1905, express the ideas then in his mind; it must be remembered that they were not written for publication, but as suggestions for his own work.

1. It is, I think, evident, from the facts of regeneration, that the theory of a nucleus as composed of specific organic determinants is hopeless. It is also evident from the behaviour of eggs, such as those of Ctenophora, that a special structure leading to determinate fate of special portions, may exist in the cell body.

2. The above facts do not invalidate conception of nuclear elements as a series of *stirps*, in Galton's sense, each containing something capable of exciting the development of any of the somatic characters, according to its position in the organism.

3. It seems necessary to regard a *stirp* as capable of exciting, not only somatic characters like those of its parents, but characters like those of its more remote ancestors under certain circumstances.

4. It is evident, from the facts of growth and regeneration, that the characters of any one stirp which become active in any one generation are determined by the position of that stirp with reference to the rest,—i.e. by a process of the same nature as Mendelian "dominance."

5. In an individual of pure race, the stirps will each contain the present and ancestral characters of that race. Assuming equality of numbers in germ cells which can fertilise each other (about which there is no real knowledge) the hybrid zygote should contain two sorts of stirps in equal numbers, each representing the race-characters of one parent. It seems possible on this assumption to develop a theory of nuclear division, which may give Mendel's results without eliminating ancestral influence,—i.e. without a theory of the "pure" gamete.

6. Such a theory would start by taking "chromomeres" as units. A chromosome of $n$ chromomeres becomes entangled in the nuclear network: say there are in the zygote $2m$ chromosomes, there will be $2mn$ chromomeres in the resting nucleus. Before division these will be gathered into $2m$ groups for an ordinary mitosis, into $m$ groups for a maturation mitosis,—the maturation division being afterwards heterotypic.

Weldon first proceeded on the basis of the non-individuality of the paternal and maternal chromosomes, and supposed a chance distribution of the chromomeres in ordinary mitosis. He investigated the problem with great wealth of numerical illustration and with a variety of hypotheses as to segregation. He concluded:

"It seems clear that such forms of segregation as I have assumed during nuclear division will not lead to a separation of zygotes into classes sufficiently sharp for Mendelian purposes, unless some persistence of effect from one mitosis to another be assumed."

He accordingly turned to the other extreme hypothesis, i.e. that from the moment of fertilisation up to the reducing division of the germ cells paternal and maternal chromosomes retain their individuality.

It is clear that this assumption disregards the reticulated stage of the chromosomes so far as its bearing on heredity is concerned. Weldon did not, however, reach the standpoint of the individual persistence and identity of the chromosomes during the reticulated state owing to a bias in favour of one or other theory. He recognised by extended numerical investigation of a number of cases that, if the chromomeres are the bearers of hereditary properties, interchange of chromomeres during ordinary mitosis was incompatible with any definite segregation in the offspring of hybrids. Throughout the theory hereinafter developed, the persistency of the individual chromosome up to the reducing division will be supposed to hold. The evidence in favour of this persistency is far from negligible, but it is hardly as complete as some Mendelians are inclined to believe. It is here accepted as a working hypothesis with the view of ascertaining whither it leads. That no hypothesis was found numerically workable and consistent with segregation phenomena, neither proves that such a hypothesis does not exist, nor demonstrates the universal truth of segregation. It merely indicates that in the present state of our knowledge we shall be most likely to reach suggestive and test results, if we follow up the idea of individuality in the chromosomes up to the reducing division.

(2) Let us suppose that each somatic cell consists of $2q$ chromosomes, effective with regard to any one inherited character*, and that each such chromosome contains $p$ determinants on which the nature of this inherited character depends. If we start with a cell belonging to an individual of pure race we may suppose these $p$ determinants alike, and the $2q$ chromosomes identical in character. We may speak of this first group of determinants as protogenic determinants and the corresponding chromosomes as protogenic chromosomes. Taking a cell belonging to an individual of a second pure race, it will also contain $2q$ chromosomes built up of $p$ determinants. These may be termed the allogenic chromosomes and allogenic determinants. Suppose $q$ protogenic chromosomes to be contained in a germ cell after the reducing division, and let them join with $q$ allogenic chromosomes in a germ cell of the second race to form a zygote. Then whatever fusion of the

* The whole number of chromosomes may be anything larger than $2q$.

nuclei may take place in the process of fertilisation, if we suppose the chromosomes to retain their individuality, there will be no interchange of chromomeres or their determinants between the allogenic and protogenic chromosomes in the zygote. Ordinary mitosis will now take place and somatic cells be indefinitely multiplied without interchange of determinants. Each hybrid cell will contain $2q$ chromosomes, $q$ with $p$ protogenic and $q$ with $p$ allogenic determinants.

The nature of the somatic characters of the hybrid will not be determinable in any way unless we have some theory as to whether the protogenic or allogenic determinants are dominant, or blend, or learn by experiment or observation that they give rise to certain characters which may differ from those of either pure race. We have a hybrid somatic cell in which there is a balance in number of the two types of determinants. As long as each ordinary mitosis divides equally the number of each kind of determinant in this heterogenic cell, it will not matter whether we suppose individual chromosomes to consist solely of determinants of one race, or each chromosome to have $p$ protogenic and $p$ allogenic determinants. The important point is that this equality of division should be maintained up to the reducing division. We shall now suppose that on the reducing division the protogenic and allogenic chromosomes fuse and interchange determinants, i.e. there is a random selection of $p$ determinants out of the total $2p$, $p$ protogenic and $p$ allogenic determinants. This is equivalent to drawing $p$ balls out of a bag containing $p$ black (= protogenic) and $p$ red (= allogenic) balls*. Or, the distribution of frequency of the several heterogenic chromosomes which contain $p$, $p-1$, $p-2$, $p-3$,... protogenic determinants is given by the successive terms of:

$$\frac{p!}{(2p)!}\left[1+p^2+\left\{\frac{p\,(p-1)}{1\,.\,2}\right\}^2+\left\{\frac{p\,(p-1)\,(p-2)}{1\,.\,2\,.\,3}\right\}^2+\ldots\right] \quad \ldots\ldots\text{(ii)}.$$

It is obvious that these terms are proportional to the squares of the binomial coefficients $(1+1)^p$.

If we take a second heterogenic cell and give it a reducing division, the frequency of the protogenic determinants in the resulting germ cells will be also given by (ii). Fertilisation of a heterogenic germ cell of $q$ chromosomes by a second will lead to the somatic cell of the offspring of the 'hybrids.' This somatic cell contains $2q$ chromosomes each containing $p$ determinants, and the frequency of the number of protogenic determinants in these chromosomes will be given by the expression

$$\left\{\frac{(p!)^2}{(2p)!}\left[1+p^2+\left\{\frac{p\,(p-1)}{1\,.\,2}\right\}^2+\left\{\frac{p\,(p-1)\,(p-2)}{1\,.\,2\,.\,3}\right\}^2+\ldots\right]\right\}^{2q} \ldots\text{(iii)};$$

* It is well known that the chances of drawing $r$, $r-1$, $r-2$, $r-3$ ... black balls out of a bag containing $u$ black and $v$ red balls are the successive terms of the series:

$$\frac{u!\,(u+v-r)!}{(u+v)!\,(u-r)!}\left\{1+r\frac{v}{u-r+1}+\frac{r\,(r-1)}{1\,.\,2}\frac{v\,(v-1)}{(u-r+1)\,(u-r+2)}\right.$$
$$\left.+\frac{r\,(r-1)\,(r-2)}{1\,.\,2\,.\,3}\frac{v\,(v-1)\,(v-2)}{(u-r+1)\,(u-r+2)\,(u-r+3)}+\ldots\ldots+\ldots\ldots\right\} \ldots\ldots\text{(i)}.$$

this is raised to the power of $2q$, since the two hybrid or heterogenic germ cells have the same constitution of chromosomes.

Now several possibilities arise at this stage of the development of the theory:

Suppose we have only a single chromosome in the germ cell, or two in the somatic cell, then $p = 1$ gives us the following condition of affairs: $\frac{1}{4}$ of the zygotes would consist of two chromosomes both with protogenic determinants, one-half of two chromosomes one with a protogenic and one with an allogenic determinant, and one-quarter with allogenic determinants alone. This is I take it pure Mendelism. The single determinant is the unit character and the protogenic and allogenic determinants are the allelomorphs. But a difficulty occurs in the fact that we must content ourselves with two chromosomes for most cases of inheritance. If we accept De Vries' view of an interchange of chromomeres bearing the determinants occurring at random between paternal and maternal chromosomes at the reducing division, then to reach pure Mendelism we must assert that only two chromosomes in the somatic cell, one in the germ cell, contain the inheritance factor for any character, and that this factor is built up of a single determinant. If we suppose for example that all the chromosomes, say $2q$ of them carry the determinant, then our distribution would be $(\frac{1}{2} + \frac{1}{2})^{2q}$, and it is clear that a homozygote would only result among the crossings of hybrids in a percentage far below the simple Mendelian 25. Thus De Vries' explanation appears to involve the narrow differentiation of the chromosomes, two only being concerned with any single character; or if more or all of them are to contain the determinant of any one character, then there must be some special mechanism to insure that all chromosomes at the random interchange of determinants before the reducing division interchange in precisely the same manner. We are not compelled to assume such a mechanism, which there is at present no cytological evidence in favour of, until it has been demonstrated that the experimental destruction, or observed absence of one or more chromosomes in a cell involves no loss of any series of characters. If a perfect zygote should develop from two germ cells, when half the nucleus of one has been removed, the problem of how De Vries' suggestion could lead to Mendelism would have to be further considered. Meanwhile let us proceed on the assumption, that there exist not a single Mendelian unit but $p$ determinants for the character either in a single chromosome of the germ cell, or in $q$ "tuned" chromosomes.

The distribution of protogenic determinants in the two, or the $2q$ "tuned," chromosomes of the somatic cells of the zygote will be given by the system of chances involved in the series

$$\frac{(p\,!)^4}{(2p\,!)^2}\left\{\pi^p + p^2\pi^{p-1}\alpha + \left(\frac{p(p-1)}{1.2}\right)^2\pi^{p-2}\alpha^2 + \frac{p(p-1)(p-2)}{1.2.3}\pi^{p-3}\alpha^3 + \ldots\right\}^2.$$

The symbols $\pi$ and $\alpha$ here refer to protogenic and allogenic determinants; the chance of $s$ protogenic and $s'$ allogenic determinants occurring in the two chromosomes ($s + s' = 2p$) or in the "tuned" $q$ pairs of chromosomes, being given by picking out the numerical coefficient in the above expansion of $\pi^s\alpha^{s'}$.

I illustrate this on the first four cases.

*Case* 1. $p=1$.

$$\frac{1}{2\times 2}(\pi+\alpha)^2 = \cdot 25\pi^2 + \cdot 50\pi\alpha + \cdot 25\alpha^2,$$

| | | |
|---|---|---|
| $\pi^2$ | 25 % | Dominance of protogenic determinants, |
| $\alpha\pi$ | 50 % | Balanced heterogenic determinants, |
| $\alpha^2$ | 25 % | Dominance of allogenic determinants. |

In this case it will be seen that the dominance of either determinant means a pure cell. The case is that of simple Mendelism and is the basis of a "pure gamete" theory.

*Case* 2. $p=2$.

$$\frac{2^4}{(4\,!)^2}\{\pi^2+4\pi\alpha+\alpha^2\}^2.$$

This gives

| | | | |
|---|---|---|---|
| $\pi^4$ | 2·78 | = 25 % | Dominance of protogenic determinant, |
| $\pi^3\alpha$ | 22·22 | | |
| $\pi^2\alpha^2$ | 50·00 | = 50 % | Balanced heterogenic determinants, |
| $\pi\alpha^3$ | 22·22 | = 25 % | Dominance of allogenic determinant. |
| $\alpha^4$ | 2·78 | | |

This case is of peculiar interest; we get absolutely the Mendelian percentages, if the somatic character follows the preponderance of a given pure race determinant, using preponderance here in a simple numerical sense. Further, the protogenic 25 per cent. would apparently breed true to the somatic character of the pure protogenic race, if the somatic character of the balanced heterozygote were, as is occasionally asserted to be the case, indistinguishable from that of the protogenic race. The peculiar suggestiveness of this result lies in the exact Mendelian properties arising on a simple view of dominance apart from any hypothesis of the pure gamete. There exists a latent allogenic determinant in the heterogenic chromosome of a large percentage of the 25 per cent. with dominant protogenic character. This, if judicious cross-breeding were adopted, might be rendered manifest in some, if only a small number, of the grandchildren of the offspring of the hybrids.

*Case* 3. $p=3$.

$$\frac{(3\,!)^4}{(6\,!)^2}(\pi^3+9\pi^2\alpha+9\pi\alpha^2+\alpha^3)^2.$$

| | | | |
|---|---|---|---|
| $\pi^6$ | 0·25 % | 29·5 % | Dominance of protogenic determinant, |
| $\pi^5\alpha$ | 4·50 % | | |
| $\pi^4\alpha^2$ | 24·75 % | | |
| $\pi^3\alpha^3$ | 41·00 % | 41 % | Balanced heterogenic determinants, |
| $\pi^2\alpha^4$ | 24·75 % | 29·5 % | Dominance of allogenic determinant. |
| $\pi\alpha^5$ | 4·50 % | | |
| $\alpha^6$ | 0·25 % | | |

Here, if the balanced heterogenic group be classed with the protogenic group, we should have the ratio 70·5 to 29·5, and most readers will remember worse experimental Mendelian quarters than this. The illustration shows that it is quite possible to get numbers approaching the Mendelian when there really exists a considerable variety in the determinantal constitution of the cells in which the protogenic race is dominant.

*Case* 4. $p=4$.

$$\frac{(4!)^4}{(8!)^2}\{\pi^4+16\pi^3\alpha+36\pi^2\alpha^2+16\pi\alpha^3+\alpha^4\}^2.$$

| | | | |
|---|---|---|---|
| $\pi^8$ | 0·020 °/₀ | $\pi^3\alpha^5$ | 24·163 °/₀ |
| $\pi^7\alpha$ | 0·653 °/₀ | $\pi^2\alpha^6$ | 6·694 °/₀ |
| $\pi^6\alpha^2$ | 6·694 °/₀ | $\pi\alpha^7$ | 0·653 °/₀ |
| $\pi^5\alpha^3$ | 24·163 °/₀ | $\alpha^8$ | 0·020 °/₀ |
| $\pi^4\alpha^4$ | 36·939 °/₀ | | |

Thus we see that there would be 31·5 °/₀ with dominance of protogenic determinant, 37 °/₀ with balanced heterogenic determinants and 31·5 °/₀ with dominance of allogenic determinant. The distribution is diverging considerably from the simple Mendelian quarter and approaching much more nearly a normal distribution (or binomial distribution) in the protogenic determinants.

No special stress is laid on the above examples, I have only taken them to indicate, how simple Mendelism passes when we increase the number of determinants concerned from a pure gamete theory to one in which we have dominance with closely Mendelian proportions, but diversity of gametic constitution in the dominant quarter, which might not cease to be latent for two or more generations. Again, on still further increasing the number of determinants, we pass to what is clearly a close approximation to the normal curve of distribution. It seems desirable, therefore, to investigate further the general theory of $p$ determinants each occurring in two or in $2q$ "tuned" chromosomes.

(3) Returning to the fundamental formula (ii), which gives the distribution of protogenic determinants in the chromosomes of the germ cell, we note the complexity of the analysis required to deal further with it in its present shape. But we see that it is formed by the squares of the terms of a symmetrical binomial. Now for $p$ even moderately large this binomial is extremely close to a normal curve, and accordingly the squares of the ordinates of this curve will represent extremely closely our series. But the squares of the ordinates of a normal curve are themselves ordinates of a normal curve. Thus the normal curve* which fits closely $(\frac{1}{2}+\frac{1}{2})^p$ is $y=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}x^2/\sigma^2}$, where $\sigma=\sqrt{\frac{1}{4}(p+1)c^2}$, and $c$ is the unit of plotting. Very frequently $\sqrt{\frac{1}{4}p}c$ is taken for $\sigma$, and $p$ being

* *Phil. Trans.* Vol. 186, A, p. 355.

large, there is no sensible difference. Since $y \times c$ gives approximately the ordinates of $\frac{1}{2^p}(1+1)^p$, we may expect

$$\frac{2^{2p}(p!)^2}{(2p)!}\, y^2 c \text{ to give the terms of (ii).}$$

Or, supposing $p$ moderately large and using Stirling's Theorem, we may anticipate that the terms of (ii) will be given fairly well by

$$Y = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{\frac{1}{8}pc^2}} e^{-\frac{x^2}{2\frac{1}{8}pc^2}} \quad \text{.....(iv).}$$

This is a normal curve of standard deviation $\sqrt{\frac{1}{8}p}\,c$, and accordingly it must again be closely given by the binomial $(\frac{1}{2}+\frac{1}{2})^{\frac{1}{2}p}$. I find that the *areas* of (iv) give (ii) closely, but that if we work with ordinates we get a better result from $\sqrt{\frac{1}{8}(p+\frac{1}{2})}\,.c$ as the standard deviation of the normal curve, i.e. neither using $p$ nor $p+1$, but the mean of these values. The following table illustrates the closeness of approximation for $p=12$:

| Accurate values from (ii) | $(\frac{1}{2}+\frac{1}{2})^{\frac{1}{2}p}$ | Normal curve: Areas ($\sigma=\sqrt{\frac{1}{2}p}=1{\cdot}22475$) | Normal curve: Ordinates [$\sigma=\sqrt{\frac{1}{2}(p+\frac{1}{2})}=1{\cdot}25$] |
|---|---|---|---|
| ·0000 | ·0000 | ·0000 | ·0000 |
| ·0001 | ·0000 | ·0001 | ·0001 |
| ·0016 | ·0000 | ·0020 | ·0019 |
| ·0179 | ·0156 | ·0185 | ·0179 |
| ·0906 | ·0937 | ·0897 | ·0887 |
| ·2320 | ·2344 | ·2306 | ·2318 |
| ·3156 | ·3125 | ·3181 | ·3191 |
| ·2320 | ·2344 | ·2306 | ·2318 |
| ·0906 | ·0937 | ·0897 | ·0887 |
| ·0179 | ·0156 | ·0185 | ·0179 |
| ·0016 | ·0000 | ·0020 | ·0019 |
| ·0001 | ·0000 | ·0001 | ·0001 |
| ·0000 | ·0000 | ·0000 | ·0000 |

It is clear that for all practical purposes the normal distribution, and even the simple binomial, may replace the expression in (ii) for the distribution of determinants in the heterogenic chromosome when $p$ is as large as 12. The errors of the actually available samples, 500 to 1000 at most of each generation, will be well within the errors of random sampling. Even if $p=4$, the agreement between (ii) and $(\frac{1}{2}+\frac{1}{2})^2$ is not wildly out:

| | |
|---|---|
| ·0143 | ·0000 |
| ·2286 | ·2500 |
| ·5143 | ·5000 |
| ·2286 | ·2500 |
| ·0143 | ·0000 |

In fact it would often be extremely difficult to distinguish between (ii) for $p=4$ and a true $\frac{1}{4}$, $\frac{1}{2}$ and $\frac{1}{4}$ system of frequency. For example, in any breeding experiment yet made we may well doubt whether it has been on a scale sufficient to distinguish between 51 and 24 per cent. as compared between 50 and 25 per cent. We are therefore justified in replacing (ii) by a normal curve for practical purposes. If we calculate areas we may take its standard deviation $\sqrt{\frac{1}{8}p}$; if we calculate ordinates we shall get a somewhat better result from $\sqrt{\frac{1}{8}(p+\frac{1}{2})}$, but for $p$ as large as 12, the results will agree within the errors of random sampling for practical series.

We have now found a fairly simple analytical form for the distribution of determinants in the chromosome of a gametic cell of the hybrid. The probability that the chromosome of the gametic cell of the hybrid will have a number of determinants of protogenic character lying between $\frac{1}{2}p+x/c$ and $\frac{1}{2}p+(x+\delta x)/c$ is

$$\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{\frac{1}{8}pc}}e^{-\frac{1}{2}x^2/\sqrt{\frac{1}{8}pc^2}}\delta x \quad \text{.....(iv) } bis.$$

The close resemblance of this result to $(\frac{1}{2}+\frac{1}{2})^{\frac{1}{2}p}$ shows that the variability of the chromosomes in the number of protogenic determinants is practically confined to the range of $\frac{1}{4}p$ to $\frac{3}{4}p$. More accurately, only two individuals per thousand would have more than $\frac{1}{2}p+\sqrt{p}$ protogenic determinants.

Now the gametic cells of both hybrid parents will have a distribution of this character, and supposing random mating among the hybrids, the frequency of protogenic determinants in the zygotes will be given by a normal curve of standard deviation $\Sigma$ where

$$\Sigma^2=\tfrac{1}{8}pc+\tfrac{1}{8}pc, \text{ or } \Sigma=\sqrt{\tfrac{1}{4}p}\,.\,c.$$

Thus the frequency with which $p+s$ protogenic and $p-s$ allogenic determinants occur in the somatic cell chromosomes of the offspring of the hybrids is given by the normal curve

$$y=\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{\frac{1}{4}pc^2}}e^{-\frac{1}{2}x^2/(\frac{1}{4}pc^2)} \quad \text{.....(v),}$$

where $s=x/c$.

(4) We have now to determine the constitution of the gametic cells of the offspring of the hybrids or to take away by the reducing division $p$ determinants at random from $p+s$ protogenic and $p-s$ allogenic determinants. The result is reached at once by substituting in (i) of p. 83 the values $u=p+s$, $v=p-s$, $r=p$. But the complexity of the result would be great (for general theoretical purposes apart from special numerical illustration) if we did not use some close-fitting continuous curve for the series. Such a curve is provided for in my memoir "On Skew Variation in Homogeneous Material*." In the notation of that memoir we find

$$\beta_1=\frac{c^2\{(1+p)^2-s^2\}}{8(p+1)}, \quad \beta_2=0, \quad \beta_3=\frac{1}{2(p+1)} \quad \text{.....(vi),}$$

* *Phil. Trans.* Vol. 186, A, p. 361.

and hence the differential equation to the curve

$$\frac{1}{y}\frac{dy}{dx} = \frac{-x}{\beta_1 + \beta_2 x + \beta_3 x^2}$$

becomes simply

$$\frac{1}{y}\frac{dy}{dx} = \frac{-2x(p+1)}{\frac{1}{4}c^2\{(1+p)^2 - s^2\} + x^2} \quad \text{.....(vii).}$$

Whence, integrating, we find

$$y = y_0 \frac{1}{[\frac{1}{4}c^2\{(1+p)^2 - s^2\} + x^2]^{p+1}} \quad \text{.....(viii).}$$

Here $x$ is measured from the modal value which is clearly the same as the mean, and corresponds to a length $\frac{1}{2}(p+s)c$, or to the number $\frac{1}{2}(p+s)$ of protogenic determinants. Further, $y_0$ will be a function of $s$ which remains to be determined.

Let $n_s$ be the total frequency corresponding to somatic cells of the offspring of the hybrids, i.e. to the cells before the reducing division. We must have

$$n_s = \int_{-\infty}^{+\infty} y\,\delta x = y_0 \int_{-\infty}^{+\infty} \frac{dx}{(\lambda_s^2 + x^2)^{p+1}},$$

where $\lambda_s^2 = \frac{1}{4}c^2\{(1+p)^2 - s^2\}$. Let $x = \lambda_s \tan\theta$, then

$$n_s = y_0 \int_{-\frac{1}{2}\pi}^{+\frac{1}{2}\pi} \cos^{2p}\theta\, d\theta/\lambda_s^{2p+1}.$$

Or

$$y_0 = n_s \lambda_s^{2p+1}/i_{2p},$$

where $i_{2p} = \int_0^{\pi} \sin^{2p}\phi\, d\phi$ is a well-known integral.

Thus (viii) becomes

$$y = \frac{n_s \lambda_s^{2p+1}/i_{2p}}{(\lambda_s^2 + x^2)^{p+1}} \quad \text{.....(ix).}$$

But $x$ is measured from $\frac{1}{2}(p+s)c$ the mean value of protogenic determinants in the gametic cells of the offspring of the hybrids. Let us write $x = -\frac{1}{2}sc + \xi$, then

$$y = \frac{n_s \lambda_s^{2p+1}/i_{2p}}{\{\frac{1}{4}c^2(1+p^2) + \xi^2 - sc\xi\}^{p+1}} \quad \text{.....(x).}$$

To find $n_s$, we must note that it is defined by the strip $y'\delta x'$ of a normal curve

$$y' = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{\frac{1}{4}pc^2}} e^{-\frac{1}{2}x'^2/(\frac{1}{4}pc^2)},$$

where $x'/c = s$, for this is the distribution for the somatic cells of the offspring of the hybrids.

Thus the frequency for the germ cells of the offspring of the hybrids of chromosomes with $\frac{1}{2}p + \xi/c$ protogenic determinants arising from offspring of hybrids with somatic cells with chromosomes with $p + x'/c$ protogenic determinants is given by the surface

$$z = z_0 \frac{\{\frac{1}{4}c^2(1+p)^2 - \frac{1}{4}x'^2\}^{p+\frac{1}{2}}}{\{\frac{1}{4}c^2(1+p)^2 + \xi^2 - x'\xi\}^{p+1}} e^{-\frac{1}{2}x'^2/(\frac{1}{4}pc^2)} \quad \text{.....(xi),}$$

where $z_0$ is a function of $p$, but not of $x'$. This is nothing else than the correlation surface between the number of protogenic determinants in somatic cell chromosomes of the offspring of the hybrids and the number of protogenic determinants in the germ cell chromosomes of the same individuals.

(xi) is again somewhat complex, but it will suffice to reduce it for most practical purposes to a normal correlation surface. Thus

$$\{\tfrac{1}{4}c^2(1+p)^2+\beta^2\}^{p+1}=\{\tfrac{1}{4}c^2(1+p)^2\}^{p+1}\left\{\left[1+\frac{\beta^2}{\tfrac{1}{4}c^2(1+p)^2}\right]^{(1+p)^2}\right\}^{1/(1+p)}$$

$$=\{\tfrac{1}{4}c^2(1+p)^2\}^{p+1}\,e^{\beta^2/\{\frac{1}{4}c^2(1+p)\}} \text{ approximately.}$$

Similarly we have

$$\{\tfrac{1}{4}c^2(1+p)^2-\tfrac{1}{4}\gamma^2\}^{p+\frac{1}{2}}=\{\tfrac{1}{4}c^2(1+p)^2\}^{p+\frac{1}{2}}\left\{\left[1-\frac{\gamma^2}{c^2(1+p)^2}\right]^{(1+p)^2}\right\}^{1/m}$$

$$=\{\tfrac{1}{4}c^2(1+p)^2\}^{p+\frac{1}{2}}\,e^{-\gamma^2/(c^2m)},$$

where $m=(1+p)^2/(\frac{1}{2}+p)$.

Thus the correlation surface, if $z_0'$ be a new function of $p$ only, is given by

$$z=z_0'\,e^{-\frac{x'^2}{c^2m}-\frac{\xi^2-x'\xi}{\frac{1}{4}c^2(p+1)}-\frac{1}{2}\frac{x'^2}{\frac{1}{4}pc^2}} \qquad \text{(xii).}$$

Writing this in the form

$$z=z_0'\,e^{-\frac{1}{2}\frac{1}{1-R^2_{s_2g_2}}\{x'^2/\Sigma^2_{s_2}-2R_{s_2g_2}x'\xi/\Sigma_{s_2}\Sigma_{g_2}+\xi^2/\Sigma^2_{g_2}\}},$$

we easily find

$$R_{s_2g_2}=\frac{1}{\sqrt{3}}\sqrt{\frac{p^2+p}{p^2+\frac{3}{2}p+\frac{2}{3}}}=\left\{\begin{array}{l}\text{Correlation between somatic and gametic}\\ \text{cells of offspring of hybrids.}\end{array}\right.$$

$$\Sigma_{g_2}=\frac{\sqrt{3}}{4}\sqrt{pc}\sqrt{\frac{p+1}{p}\,\frac{p^2+\frac{3}{2}p+\frac{2}{3}}{p^2+\frac{7}{4}p+1}}=\left\{\begin{array}{l}\text{Standard deviation of somatic}\\ \text{cells of offspring of hybrid.}\end{array}\right. \qquad \text{...(xiii).}$$

$$\Sigma_{s_2}=\tfrac{1}{2}\sqrt{pc}\,\frac{p+1}{\sqrt{p^2+\frac{7}{4}p+1}}=\left\{\begin{array}{l}\text{Standard deviation of gametic cells of}\\ \text{offspring of hybrid.}\end{array}\right.$$

Thus we find

| $p$ | $R_{s_2g_2}$ | $\Sigma_{g_2}\Big/\frac{\sqrt{3}}{4}\sqrt{pc}$ | $\Sigma_{s_2}/\frac{1}{2}\sqrt{pc}$ |
|---|---|---|---|
| 2 | ·5108 | 1·1632 | 1·0290 |
| 4 | ·5423 | 1·0865 | 1·0206 |
| 6 | ·5537 | 1·0590 | 1·0157 |
| $\infty$ | ·5773 | 1·0000 | 1·0000 |

It will be noted that as $p$ increases the correlation rapidly approaches the value ·5773 and the variability of the gametic and somatic chromosomes the values $\frac{\sqrt{3}}{4}\sqrt{pc}$ and $\frac{1}{2}\sqrt{pc}$. Thus we see that, according to the theory here developed, the chromosomes of the somatic cells are absolutely more variable in the ratio of 1 to ·866 than those of the gametic cells. But, if we deal

with their relative variabilities as measured by their coefficients of variation, the gametic cells are more variable in the ratio of 1·73 to 1. The regression of the gametic chromosomes on the somatic chromosomes gives us the value $\frac{1}{\sqrt{3}}\frac{\sqrt{3}}{4}\sqrt{p}c/(\frac{1}{2}\sqrt{p}c)=\frac{1}{2}$, or ·5 is the value this regression takes even when $p$ is a small number.

Thus, while there is no variation in the somatic cells either of the pure races or of the hybrids, and accordingly while it is impossible to find a correlation between somatic and gametic cells for these generations, its value being indeterminate, we see that as soon as we reach the offspring of the hybrids, a sensible and definite correlation exists, and this correlation has a value which may be suggestive for other forms of enquiry. The present writer is fully aware that a variety of other mathematical theories of the broad facts of cytology may be developed, and as our knowledge increases more ample theories, consonant possibly with a wider range of facts, will probably be invented. But he believes that the present theory is the first to show that there will be a definite correlation between the determinantal structure of the chromosomes in the somatic and in the gametic cells, and that its value will not improbably be found to be about ·5. Looked at from this standpoint the somatic cell precedes the germ cell of the individual, and the somatic cell might, under the proper stimulus, give rise to a germ cell. These germ cells are not of one type; they are variable, but correlated with the originating somatic cell. It is difficult, if we look at matters for the time being from this aspect, to find any basis for a "continuity of the germ plasm." Every reducing division may produce a new germ cell, and the whole group of germ cells is relatively more variable than the somatic cells from which they arise. The link between the two is one of correlation and not causation, a result produced by the random partition in the reducing division.

(5) I now turn to the somatic cells of the grandchildren of the hybrids, i.e. the children $H_2$ of the offspring $H_1$ of the hybrids $H$.

Let $p+\frac{x_1}{c}$ be the number of protogenic determinants in the chromosomes of the somatic cell of the first parental $H_1$, and $\frac{1}{2}p+\xi_1/c$ be the number of protogenic determinants in the chromosome of the gametic cell of this parent. Let $p+x_2/c$, and $\frac{1}{2}p+\xi_2/c$ be the corresponding quantities for the second parent and let us suppose the matings among the $H_1$ absolutely random. Then the number of protogenic determinants in the two chromosomes of the somatic cell of the offspring $H_2$ is $p+(\xi_1+\xi_2)/c=p+s'''$, say. We require to find the correlation surface between $s'''$ and $s'$ and $s''$ (or $x_1/c$ and $x_2/c$) or the deviations from the mean $p$ of the protogenic determinants in the offspring $H_2$ and the two parents $H_1$. Let us write $\eta=\xi_1+\xi_2$, and since when $p$ is only moderately large we have seen that $p+1$ and $m=(p+1)^2/(p+\frac{1}{2})$ may be practically replaced by $p$, we shall write for (xii)

$$z=z_0'e^{-\frac{1}{2}\left\{\frac{\xi^2}{\frac{1}{8}c^2p}-\frac{2x'\xi}{\frac{1}{4}c^2p}+\frac{x'^2}{\frac{1}{8}c^2p}\right\}}.$$

Accordingly the chance from parental somatic cells lying between $x_1$ and $x_1+\delta x_1$ and $x_2$ and $x_2+\delta x_2$ of getting gametic cells respectively between $\xi_1$ and $\xi_1+\delta\xi_1$ and $\xi_2$ and $\xi_2+\delta\xi_2$ is measured by

$$Z_0' e^{-\frac{1}{2}\frac{1}{c^2p}\{8(\xi_1^2+\xi_2^2)-8(x_1\xi_1+x_2\xi_2)+6(x_1^2+x_2^2)\}}\, dx_1 dx_2 d\xi_1 d\xi_2,$$

where $Z_0'$ is a constant depending on $p$ and $c$ only. Write for $\xi_1$, $\eta-\xi_2$, and the chance becomes

$$Z_0' e^{-\frac{1}{2}\frac{1}{c^2p}\{8(2\xi_2^2-2\eta\xi_2+\eta^2)-8(\eta x_1+(x_2-x_1)\xi_2)+6(x_1^2+x_2^2)\}}\, dx_1 dx_2 d\eta d\xi_2.$$

Now integrate this expression throughout the range of values of $\xi_2$ and we shall have the triple correlation surface connecting $x_1$, $x_2$ and $\eta$. In order to do this we must make the result a perfect square in $\xi_2$. We thus find

$$Z_0' e^{-\frac{1}{2c^2p}\{16(\xi_2-\frac{1}{2}[\eta+\frac{1}{2}(x_2-x_1)])^2+4\eta^2+5(x_1^2+x_2^2)-4\eta(x_1+x_2)+2x_1x_2\}}\, d\xi_2 d\eta dx_1 dx_2.$$

Integrating throughout the range of $\xi_2$, we only modify the constant $Z_0'$ without changing the terms in $\eta$, $x_1$, or $x_2$. Thus the frequency surface is finally of the form:

$$z = z_0 e^{-\frac{1}{2}\left\{\frac{\eta^2}{\frac{1}{4}c^2p}+\frac{x_1^2}{\frac{1}{3}c^2p}+\frac{x_2^2}{\frac{1}{3}c^2p}-\frac{2\eta x_1}{\frac{1}{2}c^2p}-\frac{2\eta x_2}{\frac{1}{2}c^2p}+\frac{2x_1x_2}{c^2p}\right\}} \quad \text{.....(xiv).}$$

Let $r$ be the correlation between chromosome in the parents' and in the offspring's somatic cells, then:

The correlation of $\eta$ with either $x_1$ or $x_2 = r$, and the correlation of $x_1$ and $x_2 = 0$.

Let $\sigma_1$ and $\sigma_2$ be the equal standard deviations of $x_1$ and $x_2$; $\sigma_3$ of $\eta$.

Then the bracket term ought to be equal to

$$\frac{1}{1-2r^2}\left\{\frac{\eta^2}{\sigma_3^2}+\frac{x_1^2(1-r^2)}{\sigma_1^2}+\frac{x_2^2(1-r^2)}{\sigma_1^2}-\frac{2\eta x_1 r}{\sigma_3\sigma_1}-\frac{2\eta x_2 r}{\sigma_3\sigma_1}+\frac{2r^2x_1x_2}{\sigma_1^2}\right\}.$$

Hence we deduce:

$$\frac{1}{\sigma_3^2}\frac{1}{1-2r^2}=\frac{4}{c^2p}, \quad \frac{1}{\sigma_1^2}\frac{1-r^2}{1-2r^2}=\frac{5}{c^2p}, \quad \frac{1}{\sigma_3\sigma_1}\frac{r}{1-2r^2}=\frac{2}{c^2p}, \quad \frac{r^2}{\sigma_1^2}\frac{1}{1-2r^2}=\frac{1}{c^2p}.$$

The second and fourth equations give us $r^2 = 1/6$, agreeing with the value obtained from the first, second and third. The first and second then give

$$\sigma_3^2 = \tfrac{3}{5}c^2p \quad \text{and} \quad \sigma_1^2 = \tfrac{1}{4}c^2p.$$

Thus finally we have:

Variability of the somatic cells in $H_1$ or in the offspring of hybrids $= \frac{1}{2}c\sqrt{p}$.

Variability of the somatic cells in $H_2$ or offspring of $H_1 = \sqrt{\frac{3}{5}}\,\frac{1}{2}c\sqrt{p}$.

Correlation of the somatic cells in parents $H_1$ and offspring $H_2 = \dfrac{1}{\sqrt{6}} = \cdot 4082$.

Regression of the filial on the parental somatic cells $= \cdot 5000$.

These results seem to be of not a little interest. They indicate that if two pure races be crossed, the variability in protogenic determinants of the chromosomes of the somatic cells will be still increasing to the third generation, i.e. it is 1·225 times as great in the $H_2$ generation as in the $H_1$ generation. The correlation,

however, is rapidly assuming a value very close to those which have been determined biometrically for large populations mating at random, and the regression has already reached that value ·5*.

It is noteworthy that the regression of the gametic cell of an individual on the somatic cell of that individual is identical in value with the regression of the somatic cell of the offspring on the somatic cell of the individual.

(6) I have not thought it worth while at present to proceed with further reducing divisions, although the results would not be without interest as determining the rate at which a stable population was established after crossing two "pure races," owing to the cessation of change in the variability of successive generations, and to the corresponding equalisation of correlation and regression. The present investigation suffices to show that we fairly rapidly approach the biometrically observed condition of things, even if we start with two pure races, and this result will be shown on another occasion to be experimentally verified. It may, of course, be said that the hypothesis which makes the inherited character depend on $p$ unit determinants in the paternal and maternal chromosomes, separated into two moieties at random in the reducing division, is not the only conceivable one. This is fully admitted, but by working it out as a first and not unreasonable hypothesis, we have reached a number of suggestive points. We see that Mendelian dominance and the Mendelian quarter may arise in cases where there is no pure gamete, and that the discovery of a latent character may need several generations of breeding. Further we see a continuous transition from simple Mendelism, through various phases of pseudo-Mendelism to distributions closely following the normal curve. We notice also that if we increase the number of determinants on which a character depends, we very soon, even if we start with two pure races, reach by hybridisation and crossing of the hybrids a population closely following the correlation found biometrically for large groups mating at random. If the hypothesis here dealt with were correct, it would follow that the Mendelians were merely working at one end of the scale, the biometricians somewhat further down. At present interesting problems which suggest themselves are: the possibility of in any way correlating the somatic and the resulting gametic cells; the existence or not of ratios actually or nearly Mendelian, but where the apparent homozygotes do not breed true; lastly, the investigation of whether or not the values of parental correlation hold closely when we have bred for only a couple of generations from the hybrids of two pure stocks. With regard to this last problem, we may hope shortly for light; Mendelian literature for the careful reader may provide answers, perhaps, to the second.

At present it does not seem needful to defend the assumptions upon which the present theory is based. It suffices that it is sufficiently wide as it stands to cover the Mendelian and the biometric views. It has been expounded here on account of its suggestiveness. What there is of good in it is Weldon's; where it may blunder is where I have failed to correctly interpret by word or symbol his ideas.

* The fact that there is in this case a difference between the correlation and regression depends upon the variability of the somatic cells of the two generations not yet being equal.

# MISCELLANEA.

## I. Some Notes on Interpolation in $n$-dimension Space.

By W. PALIN ELDERTON.

1. *Introductory.* The main idea of this paper is summed up in the principle, or the extension of the principle, that in certain cases the use of linear second differences will give a better result than the use of two-variable first differences, e.g. three points in the table chosen on a straight line may give a more accurate interpolation than the four "nearest" points. The usual formulae of interpolation assume that we require to find from tabulated values of $u_x$ a value of the function corresponding to a certain value of $x$ not among the tabulated cases. In some circumstances we have to deal with a function of two variables, and it was recently shown by Mr John Spencer* that in certain circumstances the simple one-variable interpolation formulae could be used. This saves a great deal of work in practice, and in the examples dealt with it was found to give accurate results. The object of the present paper is to supply a general method and to show that it is not necessary to limit the number of variables to two. Before proceeding to the subject itself it will be well to outline the alternative methods available. The most usual practice in two-variable interpolation is to interpolate between four values; thus if $u_{rs}$ is required and $u_{00}$, $u_{01}$, $u_{10}$ and $u_{11}$ are given, then $u_{r0}$ is found from $u_{00}$ and $u_{10}$, and $u_{r1}$ from $u_{01}$ and $u_{11}$, and then $u_{rs}$ from $u_{r0}$ and $u_{r1}$. Graphically these processes might be represented by the following diagram, in which the dots show the positions of the given values, the crosses the positions of the interpolated values, and the lines connecting the points indicate the values used in each interpolation. Of course to give the values in a figure we should have to erect perpendiculars to the surface of the paper.

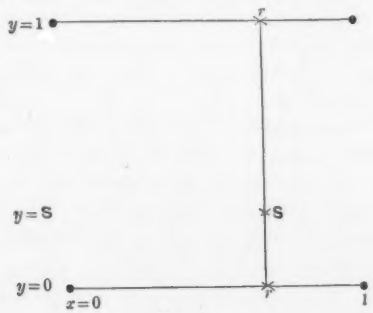



FIG. 1.

The general solution in terms of differences is obtained by expanding the right-hand side of

$$u_{rs}=(1+\Delta_x)^r\,(1+\Delta_y)^s\,u_{00},$$

* John Spencer, "Some practical hints on two-variable interpolation," *Journal of the Institute of Actuaries*, Vol. XL. pp. 293 *et seq.*

but it is, I think, better to use coefficients* instead of differences, just as one does when employing Lagrange's formula in the ordinary one-variable interpolation, and with the help of auxiliary tables such as those of Mr Herbert H. Edwards† very accurate results can be easily obtained. If however the intervals by which the given table of the function proceeds are different from those adopted by Mr Edwards, a large amount of arithmetical work is necessary. It is however worth while to recapitulate the method and increase the number of the independent variables from two to $n$. Assuming that we require to find $u_{rst\ldots(n)}$; then by Lagrange's interpolation formula we have

$$u_{rst\ldots} = r_0 u_{0st\ldots} + r_1 u_{1st\ldots} + r_2 u_{2st\ldots} \text{ etc.},$$

where $r_0$, $r_1\ldots$ are numerical coefficients depending on the number of terms and the value of $r$. And similarly

$$u_{srt\ldots} = s_0 u_{0rt\ldots} + s_1 u_{1rt\ldots} + s_2 u_{2rt\ldots},$$
$$u_{trs\ldots} = t_0 u_{0rs\ldots} + t_1 u_{0rs\ldots} + t_2 u_{0rs\ldots},$$

and so on; hence

$$u_{rst\ldots} = r_{\ldots}{}_{\ldots} t_0 \ldots u_{000\ldots} + r_0 s_0 t_1 \ldots u_{001\ldots} + \text{etc.}$$
$$= S(r_h s_k t_l \ldots u_{hkl\ldots}),$$

where $S$ stands for the summation of all combinations of $h$, $k$, $l\ldots$ for all values from 0 upwards.

The calculation of all the $r_h$, $s_k$, $t_l\ldots$ terms is done directly from Lagrange's formula and is simple though laborious, and the remainder of the work is mechanical. When the function only depends on two variables the interpolation involves fairly heavy arithmetic if only three values are given to both $h$ and $k$, while if there are three variables the work could only be completed in two or three hours, and it is therefore clear that if the ordinary one-variable interpolation formula could be used we should save a great deal of arithmetical work.

The ideas underlying the two methods can be seen graphically.

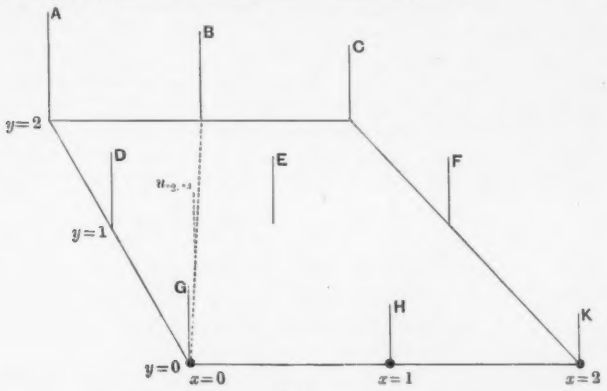



FIG. 2.

In this figure a surface $ACKG$ represents $u_{xy}$, so that when $x=0$, $y=2$, $u_{xy}$ has the value represented by the height $A$, and when $x=1$, $y=1$, $u_{xy}$ has the value represented by $E$. Now if we require to find $u_{\cdot 2, \cdot 4}$ we can approximate to the surface by the $(1+\Delta_x)(1+\Delta_y)$ method or by the Lagrange coefficient method and hence find the particular height we require, or we can choose the positions of two or more values so that the straight line passing through them also passes through the point $x=\cdot 2$, $y=\cdot 4$ and then interpolate by the one-variable formulae between the values of the function corresponding to the points on this straight line. In fig. 2 the line joining $x=0$, $y=0$ to $x=1$, $y=2$ passes through $x=\cdot 2$, $y=\cdot 4$ and by interpolation between $B$ and $G$ we can therefore approximate to $u_{\cdot 2, \cdot 4}$. The dotted lines show the process.

* W. Palin Elderton in *Biometrika*, Vol. II. Part 1.

† *Journal of the Institute of Actuaries*, Vol. XL. pp. 289—293.

2. *The Interpolation Line Method.* We may now attack the theory of this second method in greater detail.

Let us assume that from the following tabulated terms it is required to find $u_{h,k,l,\ldots}$:

$$\begin{array}{llll} u_{0,0,0,\ldots}, & u_{r,s,t,\ldots}, & u_{2r,2s,2t,\ldots}, & \text{etc.}, \\ u_{0,s,t,\ldots}, & u_{r,2s,2t,\ldots}, & u_{2r,3s,3t,\ldots}, & \text{etc.}, \\ \text{etc.}, & \text{etc.} & & \end{array}$$

This scheme is general and there is no loss of generality in assuming that the interpolated values are for integers, i.e. if the function is tabulated for $x=0, 5, 10$, etc., $y=0, 8, 16$, etc., $z=0, 2, 4$, etc., the interpolated values required would be for $x=1, 2, 3, 4$, etc., $y=1, 2, 3, \ldots$, $z=1, 3, 5, \ldots$; because if fractional values are required we can assume that the $x$'s, $y$'s, $z$'s, etc. proceed by larger differences.

Now in order to get $u_h$ we take $\frac{h}{r}$ and use the formula $(1+\Delta)^{\frac{h}{r}}u_0=u_h$, or we can take $\frac{h}{2r}$ and use $(1+\Delta')^{\frac{h}{2r}}u_0=u_h$, where $\Delta=u_r-u_0$ and $\Delta'=u_{2r}-u_0$. Again to get $u_{y+k}$ we take $(1+\Delta)^{\frac{k}{s}}u_0$ or $(1+\Delta')^{\frac{k}{2s}}u_0$ or etc., where $\Delta$ and $\Delta'$ have similar meanings.

But if $\frac{h}{r}$ be used for a two-variable, then the interpolated function found could be $u_{x+h,\, y+\frac{h}{r}s}$, and therefore in order to get terms from which interpolation can be made we must use, not $u_{00}$, $u_{r3}$, etc., but $u_{00}$, $u_{ra,s\beta}$, so that when we take $(1+\Delta_a)^{\frac{h}{ra}}u_{xy}$ we shall get $u_{hk}$. It is fairly clear that if we can fix either $ra$ or $s\beta$ the other can be obtained easily, and a little consideration of the way the $ra$ and $s\beta$ arise will show the reader that the term can be fixed by the least common multiple of $h$, $r$, $k$ and $s$.

A numerical example will make this clearer. Let $u_{0,0}$, $u_{5,8}$, $u_{10,16}$, etc. be given and assume that we require $u_{2,3}$; then the L.C.M. of 2, 3, 5, 8 being 120 we must interpolate between $u_{0,0}$, $u_{80,120}$, $u_{160,240}$, ..., for having fixed the 120, the proportion to be taken is $\frac{3}{120}=\frac{1}{40}$ and $2\times 40=80$. If a third variable is introduced the same method can be used, and if in our last numerical example we take $u_{0,0,0}$, $u_{5,8,10}$, $u_{10,16,20}$, ... as given, and require $u_{2,3,7}$, we can use $u_{0,0,0}$, $u_{240,360,840}$, $u_{480,720,1680}$, etc. The numerical work would be as follows in an imaginary example:

$$\begin{array}{rlll} & & \Delta & \Delta^2 \\ u_{0,0,0}= & 4872 & -2760 & +1656, \\ u_{240,360,840}= & 2112 & -1104, & \\ u_{480,720,1680}= & 1008. & & \end{array}$$

$$\begin{aligned} \therefore\ u_{2,3,7} &= 4872-\tfrac{1}{120}\cdot 2760-\tfrac{1}{2}\cdot\tfrac{1}{120}\cdot\tfrac{119}{120}\cdot 1656 \\ &= 4832. \end{aligned}$$

There are, of course, many objections to such a method in practice when it entails differences in the independent variable, such as 240, 360 and 840, but it can often be simplified considerably; for instance, in the last example, if we imagine the given values to extend to negative values of $x$, $y$ and $z$, we can start from $u_{5,0,10}$ and seek $u_{2,3,7}$. We notice that our fractions are now $\frac{3}{8}$, $\frac{3}{8}$, $\frac{3}{10}$ and the L.C.M. is reduced to 120. The terms to be used are $u_{5,0,10}$, $u_{-115,120,-110}$ and $u_{-235,240,-230}$, and the value of $u_{2,3,7}$ is

$$u_{5,0,10}+\tfrac{1}{30}(u_{-115,120,-110}-u_{5,0,10})-\tfrac{1}{2}\cdot\tfrac{1}{30}\cdot\tfrac{29}{30}\ \text{(second diff.)}.$$

In this way we can sometimes diminish the range of values considerably. Even here however a large table would be required for practical work, but it must be borne in mind that a three-variable table is very seldom made, and a two-variable table is frequently simplified by both the $x$'s and $y$'s proceeding by the same difference. The difficulty is however always present to some extent and it would be impossible to find values of $u$ corresponding to a number of decimal places in $x$, $y$, ..., though the method might even in these cases be of help as a rough check.

Graphically these examples can be shown very simply by the following figures. Using dots to show the position of given values, then, if the required value is $u_{2,6}$, we join (0, 0) to (2, 6) and produce it till it passes through another point $(x', y')$, then we can interpolate for $u_{2,6}$ between $u_{0,0}$, $u_{x',y'}$, $u_{2x',2y'}$, etc.

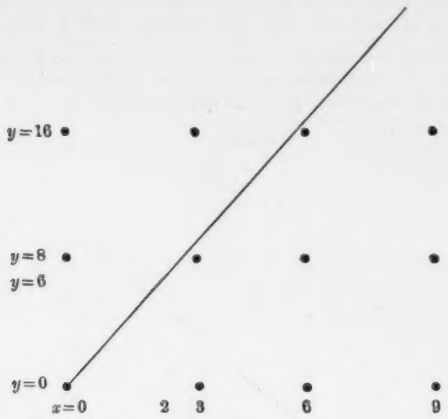



FIG. 3.

In a similar way if there are three variables the following diagram may be constructed, in which dotted lines show those parts of the figure that would not be visible if we had a solid, and the dots show the positions of given values; the "line of interpolation," as it might be called, is *OB*.

In fig. 2 we have represented the surface by giving the heights of the ordinates at the given points, and the same could have been done in fig. 3, but it would have been impossible in fig. 4

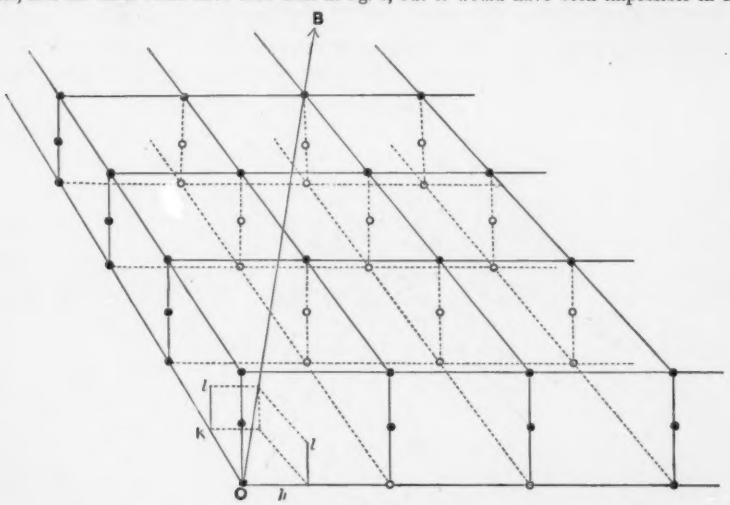



FIG. 4.

because as we are there dealing with a third-dimension body we should require a fourth dimension in order to enable us to represent the values of $u_{xy}$, as well as their positions. For the same reason we cannot give a graphical illustration of interpolation with four variables.

3. *Some particular cases.* It will probably be useful to give the "interpolation-lines" in the case of a function of two variables where the values proceed by small differences.

We shall assume that only integral values are required, and in our diagrams a heavy dot will stand for the position of a given value and a small dot for the position of a value obtained by interpolation, while the lines show the "interpolation-lines."

Difference in $x$ =3.
" " $y$ =3.

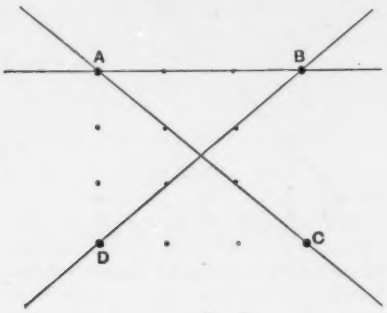



FIG. 5.

In this case the line between $AB$ is an obvious interpolation, and similar lines will be neglected in other diagrams; in fact there is only one "interpolation-line," since $AC$ and $BD$ are the same in principle; they have merely different starting-points.

Difference in $x$ =4.
" " $y$ =4.

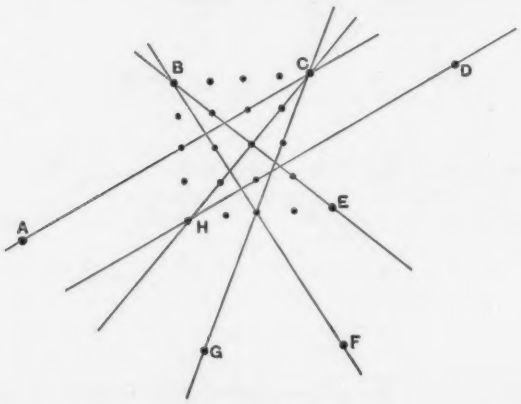



FIG. 6.

In this case there are only two "interpolation-lines," viz. *BE* and *AC*, for *CH* is the same as *BE* and *HD*, *BF* and *CG* are all the same as *AC* with different starting-points; in fact we can in general deal with any half quarter of the square *BCEH* (i.e. half of any one of the triangles formed by two diagonals and a side) and use the results as applicable to all cases.

*Example.* If we wish to interpolate for $u_{2,1}$ we should take one quarter of the way between $u_{0,0}$ and $u_{8,4}$; if a third term were required we should take either $u_{16,8}$ or $u_{-8,4}$, and the latter would generally be preferred as it is nearer the value required.

Difference in $x = 5$.
" " $y = 5$.

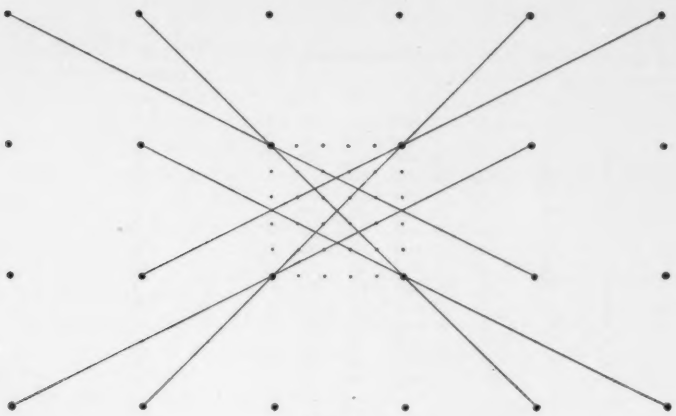


FIG. 7.

This is a very common case and there are only two necessary "interpolation-lines." It will be noticed that there is sometimes more than one line that could give the interpolation, but I have shown the one most likely to give a good result.

Difference in $x = 3$.
" " $y = 2$.

FIG. 8.

In this case in order to obtain an interpolated value for $u_{2,1}$ we require $u_{0,0}$, $u_{6,4}$ and $u_{-6,-4}$; clearly a less useful tabulation than differences of 3 in both $x$ and $y$, and not much improvement (sometimes it is worse) than differences of 4 in both $x$ and $y$. The lesson to be learnt from this is that if one can choose one's differences in tabulation it is best to take the same difference in both $x$ and $y$.

Difference in $x = 10$.

„ „ $y = 10$.

The diagram in this case becomes rather troublesome, but the following table answers the same purpose.

| $x$ | $y$ | Interpolate between | Taking $\frac{n}{10}$ of distance where $n$ is |
|---|---|---|---|
| 1 | 1 | 0, 0 and 10, 10 I | 1 |
| 1 | 2 | 0, 0 and 10, 20 II | 1 |
| 1 | 3 | 0, 0 and 10, 30 III | 1 |
| 1 | 4 | 10, 10 and −20, −10 IV | 3 |
| 1 | 5 | 0, 0 and 10, 50 V | 1 |
| 2 | 2 | 0, 0 and 10, 10 I | 2 |
| 2 | 3 | 0, 0 and 20, 30 IV | 1 |
| 2 | 4 | 0, 0 and 10, 20 II | 2 |
| 2 | 5 | 0, 0 and 20, 50 VI | 1 |
| 3 | 3 | 0, 0 and 10, 10 I | 3 |
| 3 | 4 | 0, 10 and 10, −10 II | 3 |
| 3 | 5 | 0, −10 and 10, 40 V | 3 |
| 4 | 4 | 0, 0 and 10, 10 I | 4 |
| 4 | 5 | 10, −10 and −10, 40 VI | 3 |
| 5 | 5 | 0, 0 and 10, 10 I | 5 |

This would of course be applicable where one decimal place is required and the tabulation is for integral values, but as will be noticed immediately, the number of independent "interpolation-lines" (see Roman numbers) is very great, and it is not very difficult to show that the decimal system is by no means the best method of tabulating functions. We shall however return to this after showing the application of the above diagrams to numerical work.

4. *Numerical Examples.* The following table gives the values of a function of three-variables, which will enable us to give a few examples.

| | $x=45$ | | | | $x=50$ | | | | $x=55$ | | | | $x=60$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $y$ | | | | $y$ | | | | $y$ | | | | $y$ | | | | |
| $z$ | 10 | 15 | 20 | 25 | 10 | 15 | 20 | 25 | 10 | 15 | 20 | 25 | 10 | 15 | 20 | 25 | 30 |
| 5 | 7·24 | 9·51 | 12·74 | 17·71 | 5·74 | 7·39 | 9·60 | 12·74 | 4·69 | 5·93 | 7·53 | 9·65 | 3·95 | 4·94 | 6·14 | 7·66 | 9·71 |
| 10 | 16·06 | 21·04 | 28·20 | 39·36 | 12·70 | 16·27 | 21·11 | 28·11 | 10·34 | 13·01 | 16·45 | 21·13 | 8·69 | 10·77 | 13·33 | 16·62 | 21·19 |
| 15 | 26·76 | 35·03 | 47·09 | 66·18 | 21·08 | 26·94 | 35·01 | 46·90 | 17·09 | 21·41 | 27·06 | 34·94 | 14·30 | 17·61 | 21·74 | 27·23 | 34·97 |
| 20 | 39·73 | 52·12 | 70·49 | 100·00 | 31·13 | 39·81 | 51·99 | 70·27 | 25·09 | 31·39 | 39·84 | 51·84 | 20·86 | 25·61 | 31·70 | 39·95 | 51·74 |

Let us first find a value for $x=45$, $y=17$, $z=12$; this only involves two variables, and fig. 7 tells us to interpolate between 45, 10, 5 and 45, 15, 10, etc.; hence the interpolated value required is obtained by ordinary interpolation in the following way:

| | $\Delta$ | $\Delta^2$ | $\Delta^3$ |
|---|---|---|---|
| 7·24 | 13·80 | 12·25 | 13·61 |
| 21·04 | 26·05 | 26·86 | |
| 47·09 | 52·91 | | |
| 100·00 | | | |

therefore value is

$$7{\cdot}24+1{\cdot}4\times13{\cdot}80+\frac{(1{\cdot}4)(\cdot4)}{2}12{\cdot}25-\frac{(1{\cdot}4)(\cdot4)(\cdot6)}{2.3}13{\cdot}61=29{\cdot}13,$$

while with only three terms we should have obtained 29·89. If we had used the ordinary double-interpolation method with the terms 21·04, 28·20, 47·09 and 35·03, we should have obtained 30·28, and as the true result is 29·48, we have in this case obtained a better approximation by the "interpolation-line" method. The data are not good for interpolation because the differences do not decrease, but this is no disadvantage for our present purpose.

As a second example we will take $x=60$, $y=18$, $z=11$, and we can use 14·30, 13·33 and 9·71 and obtain

$$14{\cdot}30-{\cdot}8\times{\cdot}97+\frac{(\cdot8)(\cdot2)}{2}\times2{\cdot}65=13{\cdot}73,$$

and the true value is 13·69. Two terms only would have given 13·58. With a little practice the values to be used can be easily picked out.

We will take as our final examples two three-variable cases. To obtain $x=49$, $y=18$, $z=14$ we can use 45, 20, 10 and 55, 15, 20 and take two-fifths of the difference, i.e. $28{\cdot}20+\frac{2}{5}(31{\cdot}39-28{\cdot}20)$ $=29{\cdot}48$, or we can use 50, 20, 15 and 45, 10, 10 thus getting 31·21 while if we add 55, 30, 20 which is 70·02 we get 29·93; the true value is 30·31. To obtain $x=52$, $y=14$, $z=14$ we can use 50, 10, 15 and 60, 30, 10 and obtain 21·10, or we can take 50, 15, 15 and 60, 10, 10 but the former is distinctly preferable and if we add the term 70, 50, 5 which is 17·63 it becomes practically exact.

The latter arrangement, however, shows the method at its worst. The differences run extremely awkwardly and the interpolation by first differences leads to an untrustworthy result. When using the "interpolation-line" method we are, as it were, picking out a level piece of ground on our surface before interpolating; if we choose an uneven piece of ground, i.e. if the differences are large, the result will be inaccurate.

5. *Connection of Method with the Construction of Tables.* An interesting point of some practical importance may now be examined. If we were to make a new two-variable table, what would be the best interval to use if we intended to apply the "interpolation-line" method?

We must clearly use an interval so that the terms used in the interpolations are not far apart, and in order to answer the question it becomes necessary to consider what intervals are most satisfactory from this point of view. We will first see how the various interpolations can be built up and will take the interval 5 to start with. The terms we require to find by interpolation are 1, 1; 1, 2 and 2, 2, because 1, 3 is just the same as 1, 2 but with a different starting point, for it is merely distance 1, 2 from the given term 0, 5. There are therefore only two lines necessary, one symbolised by 1, 1 and showing interpolation between terms like 0, 0 and 5, 5, and the other by 1, 2. The former gives 2, 2; we should have to take two-fifths of the distance between the 0, 0 and 5, 5 given terms instead of the one-fifth that gave 1, 1. Now let us take 7 as an interval. The terms wanted are 1, 1; 1, 2; 1, 3; 2, 2; 2, 3; 3, 3. The 1, 1 line gives 1, 1; 2, 2 and 3, 3; and the 1, 2 line gives 1, 2 and 2, 4, but the second of these is the same as 2, 3 with a different starting-point, and we only therefore require two lines. The idea is easily continued,

and it follows that if we use a prime number ($p$) as our interval, we get with each interpolation line $\frac{p-1}{2}$ cases; but as we require

$$\frac{p-1}{2}+\left(\frac{p-1}{2}-1\right)+\left(\frac{p-1}{2}-2\right)+\ldots=\frac{\left(\frac{p-1}{2}\right)\left(\frac{p-1}{2}+1\right)}{2}$$

cases in all, the number of interpolation-lines required is $\frac{1}{2}\left(\frac{(p-1)}{2}+1\right)$. This gives the following table:

| Interval | 2 | 3 | 5 | 7 | 11 | 13 | 17 | 19 |
|---|---|---|---|---|---|---|---|---|
| No. of lines | 1 | 1 | 2 | 2 | 3 | 4 | 5 | 5 |

Turning from prime numbers to those containing factors we find at once that many more lines are required. Thus we have already seen that an interval of 10 calls for no fewer than six interpolation-lines, and the intervals of 6 and 8 will both be found to require four. The interval 9 requires three lines.

Let us now compare the 11 interval with the 8 and 10 intervals and see which is the best grouping.

*Interval* 11.

| Ordinary interpolation gives | Line 1, 1 gives | Line 1, 2 gives | Line 1, 3 gives |
|---|---|---|---|
| 0, 1 | 1, 1 | 1, 2 | 1, 3 |
| 0, 2 | 2, 2 | 2, 4 | 2, 6=2, 5 |
| 0, 3 | 3, 3 | 3, 6=3, 5 | 3, 9=3, 2 |
| 0, 4 | 4, 4 | 4, 8=4, 3 | 4, 12=4, 1 |
| 0, 5 | 5, 5 | 5, 10=5, 1 | 5, 15=5, 4 |

Average interval in interpolation in $x = 8{\cdot}25$
" " " " " $y = 15{\cdot}4$
Maximum interval in ... $x = 11$
" " " ... $y = 33$

*Interval* 10.

| Ordinary interpolation gives | Line 1, 1 gives | Line 1, 2 gives | Line 1, 3 gives | Line 1, 4* gives | Line 1, 5 gives | Line 2, 5 gives |
|---|---|---|---|---|---|---|
| 0, 1 | 1, 1 | 1, 2 | 1, 3 | 1, 4 | 1, 5 | 2, 5 |
| 0, 2 | 2, 2 | 2, 4 | 2, 6=2, 4 {also by 1, 2 given above} | 2, 8=2, 2 | | |
| 0, 3 | 3, 3 | 3, 6=3, 4 | 3, 9=3, 1 | 3, 12=3, 2 | 3, 5 | 6, 5=4, 5 |
| 0, 4 | 4, 4 | 4, 8=4, 2 | | | | |
| 0, 5 | 5, 5 | which has already been obtained | | | | |

Average interpolation interval in $x = 8{\cdot}5$ or $9{\cdot}5$
" " " " $y = 23{\cdot}5$ or $22{\cdot}5$
Maximum interval in ... $x = 20$
" " " ... $y = 50$

* 2, 3 would give the same results and the averages resulting in each case are given.

*Interval* 8.

| Ordinary interpolation gives | Line 1, 1 gives | Line 1, 2 gives | Line 1, 3 gives | Line 1, 4 gives |
|---|---|---|---|---|
| 0, 1 | 1, 1 | 1, 2 | 1, 3 | 1, 4 |
| 0, 2 | 2, 2 | 2, 4 | 2, 6=2, 3 | |
| 0, 3 | 3, 3 | 3, 6=3, 2 | | |
| 0, 4 | 4, 4 | | | |

Average interval in $x = 5{\cdot}71$
" " " $y = 13{\cdot}71$
Maximum interval in $x = 8$
" " " $y = 32$

The interval 9 gives quite a good result, certainly better than 8, but considering the fewer number of values that would be required 11 seems preferable to either, and it is far better than the common decimal interval. When the number representing the interval has factors the "interpolation-lines" do not give as many values as they do when a prime number is used because, owing to the factors, there is repetition.

6. *Conclusions.* The conclusions to which these notes lead us would seem to be

(1) that interpolation can be effected by means of the ordinary one-variable formulae in $n$-dimension tables;

(2) that the method *can* give reasonably accurate results;

(3) that if the intervals can be chosen, ($a$) the same interval in all the variables will generally be better than different intervals even if the latter are rather smaller, and ($b$) the interval should be given by a prime number.

## II. Note on the Relative Variability of the Sexes in Carabus auratus, L.

By H. G. KRIBS.

(*From the Zoological Laboratory of the University of Pennsylvania.*)

The majority of biometrical investigations, which have been made regarding the comparative variability of the sexes as such, seem to have been confined to the human species*. Many observations have been made among lower organisms however, on the existence and variability of secondary sexual characters, or those morphological differentiations, apart from the sexual organs *per se*, by which the male is easily distinguished from the female of the same species. The possible significance of secondary sexual characters for evolution was first pointed out by Darwin†, as suggested by the disproportionate development of these characters in the males and females of the same species where sexual dimorphism is present, and by the difference in behaviour seemingly correlated with it. The data furnished by Darwin covered a wide range in the animal kingdom, and his conclusions have received very general acceptance.

One conclusion reached by Darwin was that in a bisexual species the female ordinarily lies closer to the morphological norm of the species than the male, and that her contribution in reproduction is inherently conservative and tends to maintain the organic stability of the norm. The male, on the other hand, was thought to vary considerably more about the species, norm.

[* The relative variability of the sexes has been dealt with in crabs, wasps, toads and a variety of other published cases, not hitherto collected together, but all tending in much the same direction as in *Carabus auratus*. Ed.]

† *Descent of Man*, Chap. VIII.

Entirely lacking in definite quantitative evidence in its favour, this general point of view regarding the relative variability of the sexes has been widely prevalent among biologists. Now it is obvious that if this generalization is a valid one the general principle involved should find expression, in some measure at least, in all sexual organisms. Whether it does or not, may be determined in a perfectly definite way by application of modern biometrical methods to the analysis of data collected by actual measurement of various characters in individuals of the two sexes.

As a small contribution toward such data, I have determined the chief variation constants of some very carefully collected material presented by Dr A. Porta *. The published data give measurements of fifteen similar characters of 84 males and 84 females of the beetle *Carabus auratus*, L. The characters measured were the following:

(1) Length of body. (2) Maximum breadth of body. (3) Length of head. (4) Width of head. (5) Length of prothorax. (6) Width of prothorax. (7) Length of antenna. (8) Length of first four joints of antenna. (9) Length of anterior leg. (10) Length of median leg. (11) Length of posterior leg. (12) Length of tarsus of anterior leg. (13) Length of mandible. (14) Length of elytra. (15) Length of first joint of antenna.

The measurements are expressed in millimetres.

In the table presented herewith, we have given, in the first column, the mean of each several character with its probable error. In the second column are likewise arranged the standard deviations from the mean of each character with their probable errors. In the third column are given the coefficients of variability and their probable errors.

From the evidence of the means of the fifteen characters measured, four—the length of posterior leg, length of anterior leg, length of tarsus of anterior leg, and total length of antenna —are so nearly equal in length that any difference is more than covered by the probable error. These characters can hardly be held to possess any secondary sexual significance. In these four characters, whose mean size is the same in both male and female, there is, furthermore, no evidence that one sex varies any more than the other. In the case of the anterior leg, the female shows a slight tendency to vary more than the male. This difference is not significant, however.

In all the other characters measured the mean is distinctly larger in the female than in the male, and we may safely assume that these differences represent a greater or less degree of secondary sexual differentiation. In four of these characters—length of first four joints of antenna, length of prothorax, length of head, and width of body—the female shows a higher index of variability than the male. In none of these, however, is the higher rate of variability of real significance, as in no case does it exceed double the probable error.

In the seven remaining characters the male shows a tendency to be more variable than the female. In four of these—length of mandible, length of first joint of antenna, length of body, and width of prothorax—the excess of variability is so slight as not to be appreciable. It is much more than offset by the size of the probable error. In the last three characters the male is probably significantly more variable than the female. In the length of elytra, length of median leg, and width of head, the excess of variability in the male is practically three times the probable error for the same.

On account of the small numbers of individuals taken in each case little emphasis can be laid upon the details of the results noted. It is worthy of consideration, however, that the four characters of equal size in each sex are equally variable, and that in the eleven remaining characters when there is a difference in size between the two sexes, only three show a greater variability in the male than in the female.

The general conclusion can only be, that, so far as the present data indicate, there is no uniform tendency for the individuals of one sex to be more variable than those of the other sex in the beetle, *Carabus auratus*, L.

* Porta, A., "Le Differenze sessuali secondarii quantitative nel *Carabus auratus*, L." *Bull. Soc. Entom. Ital. Ann.* 34, 1902.

TABLE I.

| Character | | Mean | S. D. | C. of V. |
|---|---|---|---|---|
| Length of body ... ... ... | ♂ | 22·506 ± ·077 | 1·041 ± ·054 | 4·62 ± ·24 |
| | ♀ | 22·851 ± ·081 | 1·095 ± ·057 | 4·40 ± ·23 |
| | Difference | ·345 ± ·112 | ·054 ± ·078 | ·22 ± ·33 ♂ |
| Width of body ... ... ... | ♂ | 8·994 ± ·030 | ·410 ± ·021 | 4·56 ± ·24 |
| | ♀ | 10·176 ± ·038 | ·521 ± ·027 | 5·12 ± ·27 |
| | Difference | 1·182 ± ·048 | ·111 ± ·034 | ·56 ± ·36 ♀ |
| Length of head ... ... ... | ♂ | 3·750 ± ·020 | ·270 ± ·014 | 7·20 ± ·38 |
| | ♀ | 4·036 ± ·023 | ·318 ± ·017 | 7·89 ± ·41 |
| | Difference | ·286 ± ·030 | ·048 ± ·022 | ·69 ± ·56 ♀ |
| Width of head ... ... ... | ♂ | 3·152 ± ·015 | ·199 ± ·010 | 6·30 ± ·33 |
| | ♀ | 3·396 ± ·011 | ·145 ± ·008 | 4·25 ± ·22 |
| | Difference | ·244 ± ·019 | ·054 ± ·013 | 2·05 ± ·40 ♂ |
| Length of prothorax ... ... | ♂ | 4·589 ± ·018 | ·248 ± ·013 | 5·40 ± ·28 |
| | ♀ | 4·946 ± ·021 | ·289 ± ·015 | 5·84 ± ·30 |
| | Difference | ·357 ± ·028 | ·041 ± ·020 | ·44 ± ·41 ♀ |
| Width of prothorax ... ... | ♂ | 6·693 ± ·024 | ·322 ± ·017 | 4·82 ± ·25 |
| | ♀ | 7·292 ± ·024 | ·332 ± ·017 | 4·55 ± ·24 |
| | Difference | ·599 ± ·034 | ·010 ± ·024 | ·27 ± ·35 ♂ |
| Length of elytra ... ... | ♂ | 14·411 ± ·051 | ·693 ± ·036 | 4·80 ± ·25 |
| | ♀ | 16·066 ± ·047 | ·639 ± ·033 | 3·98 ± ·21 |
| | Difference | 1·655 ± ·069 | ·054 ± ·049 | ·82 ± ·33 ♂ |
| Length of anterior leg ... | ♂ | 15·613 ± ·051 | ·698 ± ·036 | 4·47 ± ·39 |
| | ♀ | 15·589 ± ·056 | ·766 ± ·040 | 4·92 ± ·26 |
| | Difference | ·024 ± ·076 | ·068 ± ·054 | ·45 ± ·47 ♀ |
| Length of median leg ... ... | ♂ | 17·857 ± ·071 | ·965 ± ·050 | 5·40 ± ·28 |
| | ♀ | 18·304 ± ·057 | ·774 ± ·040 | 4·23 ± ·22 |
| | Difference | ·447 ± ·091 | ·191 ± ·064 | 1·17 ± ·36 ♂ |
| Length of posterior leg ... | ♂ | 24·423 ± ·081 | 1·099 ± ·057 | 4·51 ± ·24 |
| | ♀ | 24·423 ± ·080 | 1·088 ± ·057 | 4·46 ± ·23 |
| | Difference | ·000 ± ·114 | ·011 ± ·081 | ·05 ± ·33 ♂ |
| Length of tarsus of anterior leg | ♂ | 5·176 ± ·029 | 3·980 ± ·021 | 7·68 ± ·40 |
| | ♀ | 5·176 ± ·029 | 3·980 ± ·021 | 7·68 ± ·40 |
| | Difference | ·000 ± ·041 | ·000 ± ·030 | ·00 ± ·57 |
| Length of antenna ... ... | ♂ | 13·750 ± ·046 | ·629 ± ·033 | 4·58 ± ·24 |
| | ♀ | 13·685 ± ·046 | ·619 ± ·032 | 4·52 ± ·24 |
| | Difference | ·065 ± ·065 | ·010 ± ·046 | ·06 ± ·34 ♂ |
| Length of first four joints of antenna ... ... ... | ♂ | 5·329 ± ·017 | ·232 ± ·012 | 4·36 ± ·23 |
| | ♀ | 5·483 ± ·020 | ·269 ± ·014 | 4·90 ± ·26 |
| | Difference | ·154 ± ·026 | ·037 ± ·019 | ·54 ± ·35 ♀ |
| Length of first joint of antenna | ♂ | 2·063 ± ·010 | ·128 ± ·007 | 6·10 ± ·32 |
| | ♀ | 2·128 ± ·009 | ·122 ± ·006 | 5·73 ± ·30 |
| | Difference | ·065 ± ·013 | ·006 ± ·009 | ·37 ± ·44 ♂ |
| Length of mandible ... ... | ♂ | 3·304 ± ·015 | ·208 ± ·011 | 6·30 ± ·33 |
| | ♀ | 3·607 ± ·015 | ·205 ± ·011 | 5·63 ± ·29 |
| | Difference | ·303 ± ·021 | ·003 ± ·016 | ·67 ± ·44 ♂ |

## III. Variation in Flower-heads of Gaillardia Aristata.

By W. W. ROBBINS.

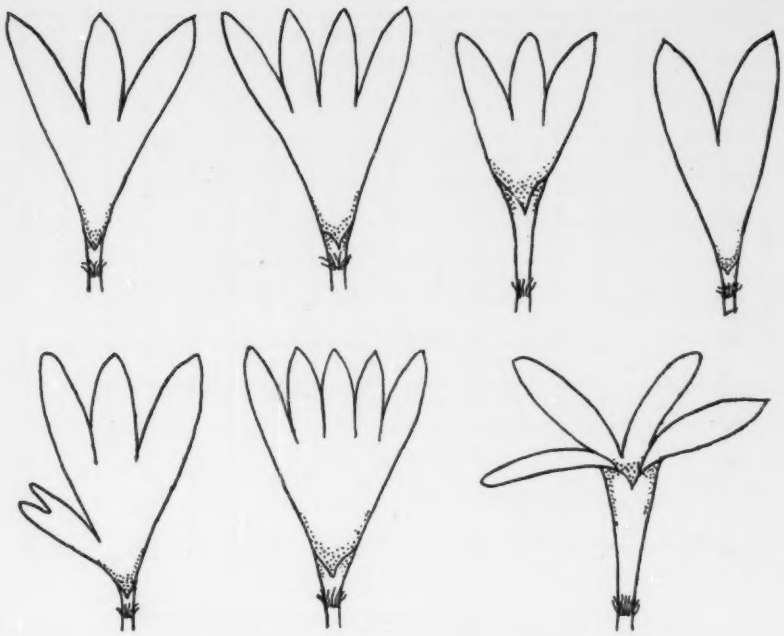
This conspicuous Composite has flower-heads about 6 cm. broad with purplish-brown centres and bright orange rays. In some heads a part of the rays are of the usual length and colour but tubular (*Torreya*, VI. 190, 1906) as was noted by the writer in a previous paper. During the summer of 1907 from July 17 to August 14 a collection of 500 flower-heads was made from localities in the neighbourhood of Boulder.

*Gaillardia aristata* commonly grows in loose soil where there is not much water. About one-half of the flower-heads studied were, however, collected at an altitude of 9000 feet, from plants growing in an aspen grove. These plants were well-developed and the rays larger than those in plants found growing in the open. Other than this no special variation has been noted in the rays of plants growing in different altitudes and habitats.

Records were made of the number and character of all the ray-flowers in each head studied. A part of the data appear in the following table.

| *1* | *2* | *3* | *4* | *5* | *6* | *7* | *8* |
|---|---|---|---|---|---|---|---|
| Number of rays in a head | Number of heads | Number of heads with all the rays ligulate | Number of heads with all the rays tubular | Number of heads with mixed ligulate and tubular rays | Percentage of heads with all the rays ligulate | Percentage of heads with all the rays tubular | Percentage of heads with mixed ligulate and tubular rays |
| *9* | 1 | 1 | 0 | 0 | 100 | 0 | 0 |
| *10* | 8 | 5 | 0 | 3 | 62 | 0 | 38 |
| *11* | 23 | 16 | 0 | 7 | 69 | 0 | 31 |
| *12* | 33 | 27 | 0 | 6 | 81 | 0 | 19 |
| *13* | 123 | 95 | 1 | 27 | 77·2 | 0·9 | 21·9 |
| *14* | 53 | 40 | 1 | 12 | 75·4 | 2 | 22·6 |
| *15* | 47 | 31 | 1 | 15 | 65·9 | 2·2 | 31·9 |
| *16* | 55 | 41 | 1 | 13 | 74·5 | 1·9 | 23·6 |
| *17* | 45 | 33 | 0 | 12 | 73 | 0 | 27 |
| *18* | 33 | 21 | 0 | 12 | 63 | 0 | 37 |
| *19* | 21 | 14 | 1 | 6 | 66·6 | 4·9 | 28·5 |
| *20* | 21 | 10 | 0 | 11 | 47 | 0 | 53 |
| *21* | 8 | 7 | 0 | 1 | 87 | 0 | 13 |
| *22* | 13 | 6 | 1 | 6 | 46·1 | 6·8 | 46·1 |
| *23* | 5 | 2 | 0 | 3 | 40 | 0 | 60 |
| *24* | 4 | 1 | 0 | 3 | 25 | 0 | 75 |
| *25* | 1 | 0 | 0 | 1 | 0 | 0 | 100 |
| *26* | 2 | 0 | 0 | 2 | 0 | 0 | 100 |
| *27* | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *28* | 3 | 0 | 0 | 3 | 0 | 0 | 100 |
| *33* | 1 | 0 | 0 | 1 | 0 | 0 | 100 |
| | 500 | 350 | 6 | 144 | 70 | 1·2 | 28·8 |

To illustrate what the table shows the 13-rayed head may be taken as an example. Columns 1 and 2 show that of the 500 heads examined 123 had 13 rays in a head. Column 3 shows that 95 of these had all ligulate rays; Column 4 shows that 1 had all tubular rays; Column 5 shows



Various forms of ray-flowers seen in the flower-heads of *Gaillardia aristata* × 1½.

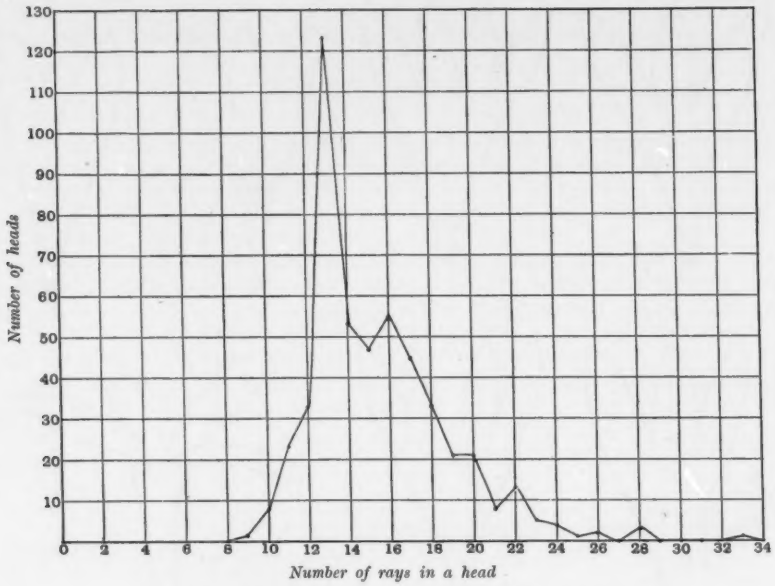
that 27 had mixed ligulate and tubular rays. Columns 6, 7 and 8 deal with percentages; it is seen that 77·2 is the percentage of 13-rayed heads with all ligulate flowers; 0·9 the percentage of 13-rayed heads with all rays tubular, and 21·9 the percentage of 13-rayed heads with mixed tubular and ligulate rays.

It is seen that when the number of rays in a head is large both ligulate and tubular rays are apt to occur. While with fewer rays in a head there are some instances in which all the flowers are tubular and many in which all are ligulate. The normal head has all ligulate rays, but a percentage of 28·8 heads with mixed ligulate and tubular rays shows a strong tendency to vary from this normal.

The accompanying curve of variation for the number of rays is computed from columns 1 and 2 of the foregoing table. It includes all rays without regard to their form. There is a well-developed mode at 13 and the usual skewness to the right which may be expected in all such studies as the present one.

The 500 flower-heads examined had a total of 7709 rays. The larger part of these rays were three-lobed and ligulate but some had as many as six lobes and a single specimen had only one lobe. In the following table the records for all of these flowers are tabulated. It includes therefore both ligulate and tubular rays.

It will be seen from the table on p. 108 that while the ligulate rays have as a rule three lobes, the tubular rays are more often four-lobed. The percentage of ligulate rays decreases and the percentage of tubular rays increases with an increase in the number of the lobes.




*Record of* 7709 *Ray-Flowers in* 500 *Flower-Heads.*

| | *f*, Ligulate | *f*, Tubular | °/₀ Ligulate | °/₀ Tubular |
|---|---|---|---|---|
| 1-lobed flowers ... | 1 | 0 | 100 | 0 |
| 2-lobed flowers ... | 193 | 0 | 100 | 0 |
| 3-lobed flowers ... | 5908 | 146 | 98 | 2 |
| 4-lobed flowers ... | 893 | 296 | 75 | 25 |
| 5-lobed flowers ... | 78 | 183 | 30 | 70 |
| 6-lobed flowers ... | 0 | 11 | 0 | 100 |
| | 7073 | 636 | 92 | 8 |

*Summary.* The normal flower-head of *Gaillardia aristata* has 13 three-lobed ligulate rays. Variation from the normal, however, is common and occurs in the number and form of the rays and in the number of lobes in a ray. The tendency to vary is toward a flower-head having a larger number of rays, mixed ligulate and tubular, the latter being four-lobed. A few heads are found with all the rays tubular.

UNIVERSITY OF COLORADO,
BOULDER, COLO.

[It is worth noting that the Fibonacci mode series (see *Biometrika*, Vol. I. p. 306), namely, 5, 8, 13, 21, 34..., is not represented in the sub-modes of *Gaillardia aristata*; and, except for the primary mode 13, there is no evidence in favour of the Fibonacci doctrine from the data. ED.]

## IV. The Effect of Errors of Observation upon the Correlation Coefficient.

By J. A. COBB.

The correlation coefficient $r=\frac{S(xy)}{N\sigma_x\sigma_y}$, where the values of $x$ and $y$ are unaffected by errors of observation. If we assume all the $x$'s and $y$'s to be subject to errors $\alpha$, $\beta$, with standard deviations $\sigma_x$, $\sigma_y$, the errors of $x$ and $y$ in every pair being independent, $\frac{S(xy)}{N\sigma_x\sigma_y}$ will become

$$\frac{S((x\pm\alpha)(y\pm\beta))}{N\sqrt{(\sigma_x^2+\sigma_{x_1}^2)(\sigma_y^2+\sigma_{y_1}^2)}}=\frac{S(xy)}{N\sqrt{(\sigma_x^2+\sigma_{x_1}^2)(\sigma_y^2+\sigma_{y_1}^2)}}=r_1.$$

So

$$\frac{r_1}{r}=\frac{\sigma_x\sigma_y}{\sqrt{(\sigma_x^2+\sigma_{x_1}^2)(\sigma_y^2+\sigma_{y_1}^2)}}.$$

In the special case where $\sigma_x=\sigma_y$ and $\sigma_{x_1}=\sigma_{y_1}$, $\frac{r_1}{r}$ becomes $\frac{\sigma_x^2}{\sigma_x^2+\sigma_{x_1}^2}$. The correctness of determinations of correlation coefficients has often been considerably affected by the neglect of this point. The question of the inheritance of sex-ratio is an interesting one in this connexion. The sex-ratio of an individual is the value which the ratio of male to total births in his fraternity would tend to assume if his fraternity were infinitely large. The standard deviation of error of sex-ratio due to taking small families is approximately $\sqrt{\frac{pq}{m}}$, where $m$ is the number of the family and $p$, $q$ the ratios of male and female respectively to the whole in all the families. The total standard deviation of the sex-ratio in families of the same size is due to the deviation $\sqrt{\frac{pq}{m}}$ and to $\sigma_x$, the variability of the sex-ratio of families. Therefore

$$\sigma_x=\sqrt{\sigma_0^2-\frac{pq}{m}},$$

where $\sigma_0$ is the total observed standard deviation. I have in this way calculated $\sigma_x$ from various data, and it turned out to be between ·03 and ·04. Now in the case of families of nine children $\frac{pq}{m}=$·028. So if the observed sex-proportion is used in forming a correlation table between the sex-ratios of parents and offspring in families of nine, the resulting correlation coefficient must be multiplied by $\frac{\sigma_x^2+\frac{pq}{m}}{\sigma_x^2}=\frac{(\cdot035)^2+\cdot028}{(\cdot035)^2}=24$ to give the true correlation between the sex-ratios of parents and offspring.

## V. On Heredity in Sex. Remarks on Mr Cobb's Note.

By KARL PEARSON, F.R.S.

Mr Cobb's point is an important one, though as a statistician I am never convinced by unpublished statistics which are said to give a certain result. But Mr Cobb's criticism would I think apply to all correlation treatment of inheritance. If we take the male offspring of a father of definite arm-length, their mean based upon four or five cases may differ very widely from the mean of all the offspring the father would have, if we supposed him capable of an indefinite number of offspring. In the same way he himself is only a single random sample from an indefinitely great array which might theoretically be attributed to his parents. I do

not see that there is anything more of the nature of "an error of observation" in taking the actual sex-ratio of the offspring as a somatic character, than in taking the actual stature or actual arm-length of the offspring. Neither represents the actual gametic character, if what is meant by that is the mean value for an indefinite number of offspring. What the biometrician is concerned with is the relation between the somatic character *as actually observed* in one generation and again in the next, and for the sex-ratio this relationship is practically zero. In the case of the sex-ratio the variability $\sigma_{x_1}$ noted by Mr Cobb is not a variability due to error of observation, but a variability characteristic of the species itself being determined by its fertility. According to Mr Cobb it is so great that it renders it impossible to determine for the sex-ratio any relation between the somatic character in parent and offspring. Is this more than a biometrician means when he says that there is no inheritance of the sex-ratio? No process, for example, of selecting families, giving many sons, would alter the sex-ratio, because the mean sex-ratio of families sprung from families with few and families with many sons is sensibly the same. I am not certain, however, that the numerical example given by Mr Cobb, i.e. $\frac{pq}{m} = \cdot 028$, is correct. For, if $p$ be really an inherited character it may take all values from 0 to 1, and accordingly we shall not get the proper value of this expression by putting $p=q=\frac{1}{2}$, the value which gives the *maximum* value of $pq$, but by putting $pq$ its actual observed mean value in man; this reduces Mr Cobb's multiplying factor to about a fifth of the value he assumes for it, and such a factor, if it were correct, would not influence the general conclusion. But Mr Heron has applied several further tests to his own data. It will be remembered that for the Whitney data in his memoir on sex inheritance Mr Heron took the sex-ratios in two generations of families having four or more members, the correlation for 2197 families was found to be $\cdot 02 \pm \cdot 01$. From the same data Mr Heron has picked out the 420 families with eight or more members, the correlation is now $\cdot 01 \pm \cdot 03$. Instead of increasing as it ought to do, on the assumption of Mr Cobb, it has got still more insignificant.

But Mr Cobb's note suggests a point which it seemed well worth while working out. If there be no correlation the resulting standard deviation of the material ought to be exactly what would be obtained on the theory of pure chance, i.e. there would be no variation due to heredity superposed on this. Mr Heron has investigated this point for 496 families given by my schedules and for 2305 families taken from the Whitney data. Suppose we take all the $n_s$ families of $s$ members; let $m_s$ be the mean number of males, $\sigma_s$ the standard deviation of this array, then $m_s/n_s$ and $\sigma_s/n_s$ will be the mean and standard deviation of the sex-ratios of this array of families of $s$ members. Let $\bar{p}$ be the mean sex-ratio of the whole series of families and $\sigma$ the standard deviation, then

$$\bar{p} = \frac{1}{N} S\left(n_s \frac{m_s}{n_s}\right) \quad \text{.....(i),}$$

$$\sigma^2 = \frac{1}{N} S\left[n_s \left\{\left(\bar{p} - \frac{m_s}{n_s}\right)^2 + \left(\frac{\sigma_s}{n_s}\right)^2\right\}\right] \quad \text{.....(ii).}$$

Now $\sigma_s/n_s$ is what Mr Cobb puts $\sigma_0$, and he makes this differ from his $pq/m$, so that he obtains a quantity $\sigma_x$, which is the variability of the sex-ratio, if the families were indefinitely great, and upon this he tells us the true heredity depends. Accordingly if $\sigma_x$ be a reality the value found for $\sigma^2$ by putting $m_s/n_s = \bar{p}$ and $\sigma_s/n_s = \bar{p}\bar{q}/n$ ought to be *less* than the actually determined value of $\sigma^2$. Mr Heron finds the following results:

| | | Value of $\sigma^2$ | | |
|---|---|---|---|---|
| Pearson's Schedules | Observed | ·1944 | Calculated | ·1946 |
| Whitney Family | „ | ·2079 | „ | ·2086 |

In both cases the calculated value is insensibly *greater* than the observed. In other words $\sigma_x$ must be imaginary, or rather within the limits of the probable error it is zero. *Shortly, if any individual had an indefinitely large family, the ratio of the sexes would not differ from that of the*

*general population, or there is no variability in the sex-ratio due to individuality. All the variation observed is actually due to the random sampling involved in taking small families.*

It appears to me therefore that Mr Cobb's note is of value, as it enables Mr Heron to confirm his results from another standpoint. For practical purposes the main problem is: Are the actual sex-ratios in parent and offspring related? The answer is: Not sensibly. Mr Cobb suggests that this result is only due to the apparent family not representing the gametic reality. This raised the question of whether any individuality at all exists for the sex-ratio. The answer appears to be none, because the standard deviation observed is precisely that which would be found were the sex-ratio for any individual that due to a random selection from the general population of a group equal in number to his family.

## VI. On the Influence of Double Selection on the Variation and Correlation of two Characters.

By KARL PEARSON, F.R.S.

In a memoir published in the *Phil. Trans.* Vol. 200 A, pp. 1—66, I have dealt with the problem of the influence of selecting $q$ characters on the variability and correlation of $n-q$ non-selected characters*. My present problem is somewhat different, and as the solution provided has been in use, and given in lectures for some years past, it may be desirable to publish it.

There are two correlated variables, 1 and 2, with standard deviations $\sigma_1$ and $\sigma_2$ and correlation $r_{12}$. In entering up a record a selection is made of the variable 1, so that its mean is shifted to $m_1$ and its standard deviation to $s_1$; let $\mu_1=s_1/\sigma_1$. In entering for the variable 2, a selection is made so that its mean is shifted to $m_2$ and its standard deviation is changed to $s_2$; let $\mu_2=s_2/\sigma_2$. As illustration suppose cards formed of characters in parent and offspring; we pass through the cards, putting a red line through every card not one of the selected parents; we pass through them again and put a blue line through every card not one of the selected offspring; the result is that cards may be thrown out because they carry a red line or a blue line or because they carry *both*. The problem is what are the standard deviations $\Sigma_1$ and $\Sigma_2$ and the correlation $R_{12}$ of the material left after this double selection. If we assume the material normal the following values are readily found by considering the probability of selecting a definite individual pair $x_1$, $x_2$ to be of the form

$$\text{const.}\times e^{-\frac{1}{2}\frac{1}{1-r_{12}^2}\left(\frac{x_1^2}{\sigma_1}-\frac{2rx_1x_2}{\sigma_1\sigma_2}+\frac{x_2^2}{\sigma_2^2}\right)}\times\frac{e^{-\frac{(x_1-m_1)^2}{2s_1^2}}}{e^{-\frac{x_1^2}{2\sigma_1^2}}}\times\frac{e^{-\frac{(x_2-m_2)^2}{2s_2^2}}}{e^{-\frac{x_2^2}{2\sigma_2^2}}}.$$

Thus the correlation surface is:

$$z=\text{const.}\times e^{-\frac{1}{2}\frac{1}{1-R_{12}^2}\left\{\frac{(x_1-M_1)^2}{\Sigma_1^2}-\frac{2R_{12}(x_1-M_1)(x_2-M_2)}{\Sigma_1\Sigma_2}+\frac{(x_2-M_2)^2}{\Sigma_2^2}\right\}}.$$

* It may be of interest to note that the whole of the formulae developed in that memoir are true far beyond the range of the Gaussian distributions for which they were proved. They are universally true provided we start with a generalised notion of correlation as involving the maximum dependence of one variable on an arbitrary linear function of $(n-1)$ other variables.

Hence we deduce:

$$\Sigma_1^2 = \sigma_1^2 \frac{\mu_1^2\{1-r_{12}^2(1-\mu_2^2)\}}{1-r_{12}^2(1-\mu_1^2)(1-\mu_2^2)} \quad \text{.....(i)}$$

$$\Sigma_2^2 = \sigma_2^2 \frac{\mu_2^2\{1-r_{12}^2(1-\mu_1^2)\}}{1-r_{12}^2(1-\mu_1^2)(1-\mu_2^2)} \quad \text{.....(ii),}$$

$$R_{12} = r_{12} \frac{\mu_1\mu_2}{\sqrt{1-r_{12}^2(1-\mu_1^2)}\sqrt{1-r_{12}^2(1-\mu_2^2)}} \quad \text{.....(iii),}$$

$$\frac{M_1}{\sigma_2} = \frac{m_1}{\sigma_1} \frac{1-r_{12}^2(1-\mu_2^2)}{1-r_{12}^2(1-\mu_1^2)(1-\mu_2^2)} + \frac{m_2}{\sigma_2} \frac{\mu_1^2 r_{12}}{1-r_{12}^2(1-\mu_1^2)(1-\mu_2^2)} \quad \text{.....(iv),}$$

$$\frac{M_2}{\sigma_2} = \frac{m_1}{\sigma_1} \frac{\mu_2^2 r_{12}}{1-r_{12}^2(1-\mu_1^2)(1-\mu_2^2)} + \frac{m_2}{\sigma_2} \frac{1-r_{12}^2(1-\mu_1^2)}{1-r_{12}^2(1-\mu_1^2)(1-\mu_2^2)} \quad \text{.....(v).}$$

These formulae will be found useful (especially in deducing $r_{12}$ from $R_{12}$) in records in which there has been independent selection of two related individuals, without regard to their relationship.

## VII. On certain points concerning the Probable Error of the Standard Deviation*.

By RAYMOND PEARL.

The purpose of this paper is to discuss two problems of considerable practical importance in all biometrical investigations. These problems presented themselves in acute form in some studies of fecundity which the writer has at present under way. It was decided to be necessary to get definite answers to them before going farther with the work mentioned. In the belief that the matter is of general interest to workers in biometry it is presented here. The two points may be stated as follows:

I. It has been shown† that if in any frequency distribution $\sigma_{\mu_q}$ be the standard deviation for errors in the $q$th moment coefficient $\mu_q$, taken about the mean, and $\sigma$ be the standard deviation of the distribution, then

$$\sigma_{\mu_q} = \sqrt{\frac{\mu_{2q} - \mu^2{}_q - 2\mu_{q+1}\mu_{q-1} + q^2\sigma^2\mu^2{}_{q-1}}{n}},$$

where $n$ denotes the number in the sample. In this expression put $q=2$, and then, since $\mu_1=0$, we have at once

$$\text{Probable error of } \mu_2 = \cdot 67449\sqrt{\frac{\mu_4 - \mu_2^2}{n}}.$$

Further since $\sigma=\sqrt{\mu_2}$, we have

$$\text{P.E. of } \sigma = \cdot 67449\sqrt{\frac{\mu_4 - \mu_2^2}{4\mu_2}\Big/ n} \quad \text{.....(i).}$$

This is the true value of the probable error designated, whatever be the type of the frequency curves. But, for the normal curve, since there $\mu_4 = 3\mu_2^2$, (i) reduces at once to

$$\text{P.E. of } \sigma = \cdot 67449\sqrt{\frac{1}{2n}}\,\sigma,$$

or as it is usually written

$$= \cdot 67449\frac{\sigma}{\sqrt{2n}} \quad \text{.....(ii).}$$

* Papers from the Biological Laboratory of the Maine Agricultural Experiment Station, No. 1.

† *Biometrika*, Vol. II. p. 276.

Now this value (ii) is the one almost universally used in calculating the probable error of the standard deviation, quite without regard to the type of the distribution under discussion. Indeed one suspects, from statements made in memoirs and handbooks of biometrical methods, that the impression prevails among biologists that (ii) is the true and unique expression for this probable error in every case. Obviously (i) will not equal (ii) except in the special case of the normal curve. In the common use of (ii) for (i) it is, of course, assumed, consciously or unconsciously, that the deviation of the curve from normality is not such as to cause, for *practical* purposes, any significant error as far as these probable errors are concerned. So far as the writer is aware no one has attempted to test in a concrete way the validity of this assumption in specific instances. Yet the matter is obviously a very important one. Probable errors lie at the very basis of all biometrical reasoning. If in certain cases the probable errors are likely to be considerably in error when calculated in a certain way, any reasoning based upon them will have a most insecure foundation.

These considerations suggest the following practical problem to the working biometrician: When a frequency curve is *not* a normal curve, is the error made in the probable error of the standard deviation by assuming normality (using (ii) for (i)) ever likely to be great enough to be of *practical* significance in drawing conclusions from data? This is our first problem.

II. It has recently been pointed out by Sheppard* that in calculating the probable errors of the mean and standard deviation it is theoretically not justifiable to use the moments to which his corrections have been applied. Instead the "raw" moments should be used. The theory of this is stated as follows by Sheppard (*loc. cit.*), $\mu$ being used to denote corrected and $\pi$ to denote raw moments: "Similarly if we take $\mu_2$ to be equal to $\pi_2' - \frac{1}{12}h^2$, where $\pi_2'$ is the value of $\pi_2$ as calculated from the actual measurements, the error in $\mu_2$, due to the limitation of number of observations, is equal to the error in $\pi_2$; and the mean square of this latter error is $(\pi_4 - \pi_2^2)/n$. Hence the probable error in the standard deviation will be, not

$$\cdot 67449 \sqrt{\frac{\mu_4 - \mu_2^2}{4\mu_2}} \Big/ n \quad \text{.....(iii),}$$

but

$$\cdot 67449 \sqrt{\frac{\pi_4 - \pi_2^2}{4\mu_2}} \Big/ n \quad \text{.....(iv).}"$$

Now while it is clear that *theoretically* the uncorrected moments should be used, will it *practically* make any significant difference in the results if the probable errors of the mean and standard deviation are calculated from the corrected moments? This is our second problem.

To get a "working" answer to these two questions it was decided to calculate for a series of curves the probable error of the standard deviation by each of the different methods possible. These possibilities are, of course, first to calculate the probable error by (i) with (*a*) corrected and (*b*) "raw" moments, and then by (ii) using again the corrected and "raw" values of $\sigma$ successively. With these results in hand for a number of curves we shall be able to form a practical judgment of the necessity or desirability of using the more refined methods in ordinary biometrical investigations. In choosing curves on which to make these tests it was decided to take a rather extreme example of each of Pearson's six types of curves. We may thus expect to get some idea of the maximum effect produced by calculating the probable error under discussion in the several different ways.

The curves actually chosen for the work were the following:

*Type I.* Curve for variation in leaf number of whorls on "all branches" of certain *Ceratophyllum* plants†. This is an extreme example of a skew curve of Type I. It is figured on p. 34 of the memoir cited.

* *Biometrika*, Vol. v. p. 455, 1907.

† Pearl, R., Pepper, O. M., and Hagle, F. J. "Variation and Differentiation in Ceratophyllum." Carnegie Institution of Washington, *Publication* No. 58, p. 32.

*Type II.* A curve from some unpublished material in the hands of the writer.

*Type III.* Curve of frequency of barometric heights, Churchstoke station*. This curve is figured as Fig. V, of Plate II, of the memoir cited.

*Type IV.* Curve of variation in length of the carpopodite of leg III in the crayfish *Cambarus propinquus* Girard†.

*Type V.* Curve of variation in the number of lips in the medusa, *P. pentata*‡. The curve is figured on p. 456 of the memoir cited.

*Type VI.* Curve of variation in leaf number in main stem whorls of *Ceratophyllum*§. The curve is figured on p. 34 of the memoir cited.

Certain of the constants of these curves are given in Table I. From the constants it will be seen that all these curves differ widely from the normal or Gaussian curve. Taken as a whole they are fairly representative of the conditions which one finds in frequency curves in practice when large numbers of individuals are dealt with. It will be noted that in some of these curves as tabled, Sheppard's corrections of the moments have been used, while in other cases they have not.

TABLE I. *Constants of Curves.*

| Constant | Curve of Type I | Curve of Type II | Curve of Type III | Curve of Type IV | Curve of Type V | Curve of Type VI |
|---|---|---|---|---|---|---|
| N | 1954 | 275 | 4018 | 283 | 996 | 374 |
| Unit of grouping | 1 leaf | 15 | 1/10 in. | 3/10 mm. | 1 | 1 leaf |
| $\mu_2$ | 1·5350 | 6·3761 | 12·6642 | 7·2794 | ·3090 | ·8798 |
| $\mu_4$ | 5·8052 | 111·2316 | 511·4453 | 246·7453 | 1·1817 | 4·0691 |
| $\beta_1$ | ·2112 | ·0008 | ·1258 | ·7276 | 4·1683 | 1·1928 |
| $\beta_2$ | 2·4637 | 2·7360 | 3·1889 | 4·6565 | 12·3760 | 5·2568 |
| $\kappa_1$ | −1·7060 | − ·5303 | + ·0005 | + 1·1302 | + 6·2469 | + ·9354 |
| $\kappa_3$ | − ·1002 | − ·0011 | +198·5780 | + ·5738 | + 1·0659 | +1·2455 |
| $\sigma$‖ | 1·2390 | 37·8765 | ·3559 | ·809 | ·5559 | ·9380 |
| Skewness | − ·6119 | − ·0171 | + ·1773 | + ·3293 | − ·5294 | − ·4270 |
| Sheppard's corrections | Not used | Used | ¶ | Used | Not used | Not used |

We may turn now to the consideration of our first problem. Taking the values of the moments given in Table I for the six curves I have calculated the probable error of the standard deviation for each curve, first according to equation (i) above, and then according to equation (ii). It will be noted that this procedure takes no account of whether Sheppard's corrections have been applied to the moments of the several distributions. It is here assumed that (i) is the absolutely true formula in every case. The results are shown in Table II. This table further gives the absolute differences for each pair of probable errors calculated in the two ways. Finally an expression of the relative magnitude of the error made by the use of the approximate formula was obtained by finding in each case what percentage of the true probable error the

* Pearson, K., and Lee, A. *Phil. Trans.* Vol. 190, A, pp. 423—469, 1897.

† Pearl, R., and Clawson, A. B. "Variation and Correlation in the Crayfish." Carnegie Institution of Washington, *Publication* No. 64, p. 7, Table I.

‡ Pearson, K., *Phil. Trans.* Vol. 197, A, p. 455, 1901.

§ Pearl, Pepper, and Hagle. *Loc. cit.* p. 32.

‖ Given throughout in concrete units, not in units of grouping.

¶ The moments in this case were corrected by considering the frequency areas as trapezia. Cf. Pearson, *Phil. Trans.* Vol. 186, A, p. 350.

approximate one was and then subtracting the percentage from 100. The figures so obtained are given in the last line of the table opposite the entry "Relative Error."

TABLE II.

*Probable Errors of Standard Deviations.*

| P. E. of $\sigma$ | Curve of Type I | Curve of Type II | Curve of Type III | Curve of Type IV | Curve of Type V | Curve of Type VI |
|---|---|---|---|---|---|---|
| Calculated by (i) ... | ·0114 | 1·0149 | ·0028 | ·0310 | ·0200 | ·0337 |
| Calculated by (ii) ... | ·0134 | 1·0893 | ·0027 | ·0229 | ·0084 | ·0231 |
| Absolute Difference ... | −·0019 | −·0644 | +·0001 | +·0081 | +·0116 | +·0106 |
| Relative Error ... | 17 % | 7 % | 4 % | 26 % | 58 % | 31 % |

From this table we note at once the following facts:

(*a*) As was to be expected the approximate formula for P.E. of $\sigma$ in the case of these skew curves never gives the true value.

(*b*) The probable error obtained by the approximate formula may be either in excess or defect of the true value.

(*c*) The amount of the deviation of the approximate from the true value varies in the different curves but may be very considerable. Thus in the case of Type V curve the probable error from the approximate formula is less than half as large as it should be. In the Type VI curve the approximate value is but 69 % of the true, and the Type IV curve 74 %. It will be noted that not only are these deviations of the approximate from the true values large in amount, but further they are in the direction most likely to cause serious error in practical biometrical work. It is not so serious a matter when an approximate formula makes a probable error too large as when it makes it too small.

(*d*) In the case of the curves of Types I, II and III the deviations of approximate from true values are all small, and in the first two cases the approximate value is in excess of the true. So far as we may judge from the curves here discussed, it would appear that the most serious difficulty from the use of the approximate probable error formula is to be expected in curves of Types IV, V and VI. Thinking that possibly the great deviations observed in the case of Type V and VI curves might be exceptional and due to the particular examples chosen, I took another Type VI curve and calculated the P.E. of $\sigma$ as before. This curve is one discussed by Pearson (*Phil. Trans.* Vol. 197, A, p. 451). It gives the variation in age of the brides in 28,454 Italian marriages. The values of the moments indicate a curve of Type VI but the criterion $\kappa_2$ is so nearly 1 that a Type V curve gives a good graduation. The results of the probable error calculations (using corrected values of the moments) are shown in the following table:

| P. E. of $\sigma$ | Curve of Type VI |
|---|---|
| Calculated by (i) ... | ·0177 |
| Calculated by (ii) ... | ·0103 |
| Absolute Difference ... | +·0074 |
| Relative Error ... | 42 % |

We see here just as before that the approximate value is largely in defect of the true value.

(*e*) Considering Table II in connection with Table I it is evident that there is no simple relation between what I have called the "Relative Error" and any of the constants of the curves. That is to say, even though we know the ordinary constants of a given curve it will not be possible to say beforehand whether the approximate formula for the P.E. of $\sigma$ will give a reasonably good value in that case.

Taking into account all these facts I think there can be no doubt regarding the answer to our first question. The answer definitely is that when a distribution is not closely normal, there *is* a considerable likelihood that the assumption of normality made in calculating the probable error of the standard deviation by the approximate formula may introduce an error great enough to be of practical importance in drawing conclusions. Whenever a piece of biometric work involves the comparison of standard deviations to determine whether they are significantly different among themselves great caution should be exercised about making conclusions depend on probable errors calculated by the approximate formula.

We may now turn to the consideration of our second problem (cf. p. 113 *supra*). In attempting to get at an answer to it the plan was to calculate for each of the curves given in Table I (with the exception of the Type III curve) the probable error of $\sigma$ first according to equation (iv) and then according to equation (iii), using, of course, the proper values of $\mu$ and $\pi$. The same calculations were also made for the curve obtained from the Italian marriage statistics (Type VI) cited above. The results of these computations are set forth in Table III. The "Relative Errors" in this case are obtained as before by taking the percentage which the approximate is of the absolutely true value and subtracting this percentage from 100.

## TABLE III.

*Probable Errors of Standard Deviations. Influence of Sheppard's Correction.*

| | Curve of Type I | Curve of Type II | Curve of Type IV | Curve of Type V | Curve of Type VI | Curve of Type VI (Italian Marriages) |
|---|---|---|---|---|---|---|
| $\cdot 67449\sqrt{\frac{\pi_4-\pi_2^2}{4\mu_2}}\Big/n$ | ·0118 | 1·0301 | ·0312 | ·0234 | ·0355 | ·01775 |
| $\cdot 67449\sqrt{\frac{\mu_4-\mu_2^2}{4\mu_2}}\Big/n$ | ·0108 | 1·0149 | ·0310 | ·0226 | ·0340 | ·01771 |
| Difference ... | +·0009 | +·0152 | +·0002 | +·0009 | +·0015 | +·00004 |
| Relative Error ... | 7·8 °/₀ | 1·5 °/₀ | 0·6 °/₀ | 3·8 °/₀ | 4·2 °/₀ | 0·2 °/₀ |

From this table we note:

(*a*) That the absolute differences resulting from calculating the probable errors in the two different ways are in every case very small. In only one case (the Type II curve) is the absolute difference greater than 1 in the third decimal place. Now it is, of course, obvious that in ordinary statistical work the probable errors will very rarely be tabled to more than the third place of figures. It would appear then, from the absolute differences here, that the error made by using the approximate formula (iii) instead of (iv) is for *practical* purposes insignificant.

(*b*) Relative to the magnitude of the probable errors themselves the differences are very small. It is seen from the last line of the table that approximately 8°/₀ is the maximum relative error made and in the majority of cases the relative error is much smaller than this.

From these results the answer to our second question is, I think, clear, and may be stated in the following way: it is *not* a matter of practical significance in ordinary statistical work that the raw moments be used to calculate the probable error of a standard deviation which has itself been deduced from a moment to which Sheppard's correction has been applied. The error made by using equation (iii) instead of equation (iv) is not likely in the great bulk of ordinary biometrical work ever to reach the magnitude where it will be of any *practical* consequence whatever. One should, of course, always keep in mind the existence of this source of error, and if very fine, close work is to be done take pains to eliminate it. It is obvious, however, that the occasions where it will have to be taken into account will be extremely rare.

Putting all our results together we may summarily state the practical conclusion as follows:

In calculating the probable error of the standard deviation of frequency distributions which deviate sensibly from the normal curve, great caution must be exercised in using the customarily applied formula for this probable error $\left(\cdot 67449 \frac{\sigma}{\sqrt{2n}}\right)$. The use of this approximate formula is not unlikely to lead to serious error when standard deviations are compared on the basis of their probable errors. It will be safest, in the absence of special investigation of the point, for each curve or group of curves under treatment, to calculate the probable error of $\sigma$ by the formula $\cdot 67449 \sqrt{\frac{\mu_4 - \mu_2^2}{4\mu_2}} \Big/ n$. In using this formula for the probable error it is practically immaterial whether the moments used have or have not been corrected according to Sheppard's formulae.

[*Note.* The standard deviation of the standard deviation $\sigma$ is known to be $\frac{\sigma}{\sqrt{2n}}\sqrt{1+\frac{1}{2}\eta}$, where $\eta$ is the kurtosis $=\beta_2 - 3$. For $\eta$ small the multiplying factor is $1+\frac{1}{4}\eta$, and supposing we may neglect for practical purposes an error of 5% in a probable error, we conclude that the ordinary formula may be applied as long as the kurtosis lies between $\pm \cdot 2$ or for $\beta_2 = 2 \cdot 8$ to $3 \cdot 2$. This agrees with Dr Pearl's Table II, where only the curve of Type III has its kurtosis within this range. Thus the kurtosis alone, independently of the type, settles whether the ordinary formula is sufficient or not. A further point not yet discussed but of some importance is the significance of the "probable error" at all when we are taking the standard deviation of standard deviations in the case of very skew material. How far does the distribution of standard deviations follow a normal curve? The answer depends largely, I think, on the size of the sample, and the deduction of conclusions from the "probable error" of $\sigma$ may lead to greater error than even neglecting $\eta$, unless we are fairly certain that our $\sigma$ follows a normal distribution. See the paper by "Student" in the present number of this Journal. K. P.]

## VIII. Addendum to Memoir: "Split-Hand and Split-Foot Deformities."

*Biometrika*, Vol. VI. pp. 26—58.

Since the final proof sheets of the above paper were sent to press, several papers have appeared to which reference is necessary.

The evidence of the Bateson School to show that Mendelism is applicable to deformities in man was criticised in the memoir. At that time stress had been laid on three family trees, that of Farabee, an instance of hypophalangia, and those of Nettleship, instances of congenital cataract. The tendency for the deformity to abate in successive generations was noted, for in these families the last generation in each is the only one which shows the undeformed out-

numbering the deformed. The criticism was justified in that the three families had been repeatedly made use of to support the view criticised. Another family of "brachydactyly" is now instanced by Punnett (*Proc. Roy. Soc. of Med.* Feb. 28, 1908), who states that "no member of a brachydactylous family who is free from the defect can transmit it to his or her offspring." For certain families this conclusion might be justified from a consideration of the facts, and apart from the question of Mendelism. But that it is of general applicability is disproved by the family recorded by Hasselwander (*Zeitschr. f. Morph. u. Anthrop.* Bd. 6, 1903, S. 510—526), in which such transmission actually occurs. Drinkwater in the article referred to by Punnett (*Proc. Roy. Soc. Edin.* Vol. 28, 1908, p. 38), states, that of 75 descendants of abnormal parents in the family he records, 39 were abnormal, "a result corresponding with what we should expect from Mendel's law." The statement is misleading, for it implies that 36 of the offspring were normal. As a matter of fact four of these were uncertain, and there is as much reason to consider the ratio 43 to 32, as 39 to 36. The mean, 41 to 34, is the fairest expression of the ratio, and it shows the hyperdominance of the deformed, common to so many malformations.

In regard to the possibility of stability of a deformity, the Nettleship family is of interest (*Ophthalm. Soc. Trans.* Vol. 27, p. 269), in that "night-blindness" has been traced through nine generations and shows no marked tendency to abate. A single instance of this character, though suggestive, does not constitute proof; and in any case the conclusion could be applied solely to "night-blindness."

The point which requires emphasis is that a statement to the effect that a family of brachydactyly is an example of Mendelian heredity (Punnett, *loc. cit.*), or the inclusion of that deformity in lists which purport to show examples of instances in which Mendelism has been witnessed (Bateson, *Progress. Rei Botan.* 1906), is not justified by present evidence*. It may be, and very probably is, the case, that Mendelism applies to certain hereditary human deformities; but the conclusions which are being drawn, or implied, conclusions having a serious sociological aspect, are at present ahead of the facts at our disposal.

THOMAS LEWIS.

## IX. On a Formula for Determining $\Gamma(x+1)$.

EDITORIAL.

The statistician so often requires the value of the $\Gamma$-function, or more usually the value of $\Gamma(x+1)/(x^x e^{-x})$, that a suggestion as to a ready means of determining it may not be out of place. Every biometrician may be supposed to have at hand Barlow's tables and an arithmometer of some form. The usual series, Stirling's or the Bernoulli number formula, involve inverse powers of $x$, and thus do not lend themselves to rapid calculation in cases where $x$ may involve three or four figures. Forsyth's formula, i.e.

$$\Gamma(x+1)=\sqrt{2\pi}\left\{\frac{\sqrt{x^2+x+\frac{1}{6}}}{e}\right\}^{x+\frac{1}{2}} \quad \text{.....(i),}$$

is very accurate. It may be read for our purposes as

$$\log\frac{\Gamma(x+1)}{x^x e^{-x}}=0{\cdot}181,9427+\tfrac{1}{2}(x+{\cdot}5)\log\{(x+{\cdot}5)^2-\tfrac{1}{12}\}-x\log x \text{ ..... (ii).}$$

* Those interested should refer to Joachimsthal (*Virch. Archiv*, 1898, Bd. 6, S. 429). It is obvious from his paper that so-called hypophalangia or brachydactylia is a by no means simple malformation, and is possibly not an entity at all. The number of digits affected and the grade of affection is open to wide variation. There is an instance in the paper (Fall 1), of a symmetrical affection of two fingers, transmitted through a normal individual. Ammon (*loc. cit.* S. 96), quotes the Kellie family, in which hypophalangia is said to have passed through ten generations, and shown sex limitation in transmission.

In this case a square has to be formed, a difference taken, two logarithms looked out, and two multiplications by separate factors undertaken. It is an advantage for arithmometer work that our multipliers should remain the same, especially if a series of values of the function have to be tabled. I accordingly suggest the following formula

$$\log \frac{\Gamma(x+1)}{x^x e^{-x}} = 0{\cdot}399{,}0899 + \tfrac{1}{2}\log x + {\cdot}080{,}929 \sin \frac{25°{\cdot}623}{x} \quad \text{.....(iii).}$$

The advantages of this formula are: (i) the multiplying factors are constants, (ii) the factor $\frac{1}{2}$ can be applied in taking out the logarithm, (iii) the reciprocals of $x$ are tabled, whereas the inverse powers are not, (iv) all the terms are to be added, which saves some liability to error.

I have slightly adjusted the constants of the sine term, so that the multiplying factors are fairly simple and for low values of $x$ tend to allow for the neglected $\frac{1}{x^5}$ term. The calculation requires: one use of Barlow's tables, one use of a logarithmic table, and one of a table of natural sines, with two applications of the arithmometer.

As illustration take $x = 7{\cdot}75$.

Barlow: $\frac{1}{x} = {\cdot}129{,}0323$.

Arithmometer: ${\cdot}129{,}0323 \times 23°{\cdot}623 = 3°{\cdot}306195 = 3°\ 18'{\cdot}3717$ at sight.

Table of Natural Sines: $\sin 3°\ 18'{\cdot}3717 = {\cdot}057{,}6719$.

Arithmometer: ${\cdot}080{,}929 \times {\cdot}057{,}6719 = {\cdot}004{,}6673$.

Accordingly, $\log \frac{\Gamma(x+1)}{x^x e^{-x}} = \left\{ \begin{matrix} {\cdot}399{,}0899 \\ {\cdot}444{,}6509 \\ {\cdot}004{,}6673 \end{matrix} \right\} = {\cdot}848{,}4081.$

Since $\log x^x e^{-x} = {\cdot}848{,}4081$, we have $\log \Gamma(x+1) = 4{\cdot}374{,}7139$.

Working with 7-figure logarithms only, and using the reduction formula and Legendre's Tables, I find $\log \Gamma(x+1) = 4{\cdot}374{,}7140$, which agrees absolutely within the limits of error of the 7-figure tables.

For this particular value of $x$, Formula (ii) gives $\log \Gamma(x+1) = 4{\cdot}374{,}7105$, not quite so good an agreement.

Generally Formula (iii) is in error in the value of $\Gamma(x+1)$ 1 in 50,000 for $x=2$, 1 in 900,000 for $x=6$, and correct to 7 figures if $x=10$. It is thus more than sufficient for all statistical purposes, where 1 in 1000 is quite a permissible difference with the usual values of the probable errors. For the case of $x=2$, the worst possible: $2! = \Gamma(3)$ is given as 1·99996 instead of 2. Formula (i) gives 1·99948.

We may, I think, safely conclude that Formula (iii) gives the $\Gamma$-function throughout the whole range of the variable * with an adequacy more than sufficient for statistical purposes, and with extremely little numerical labour.

* Of course Legendre's Tables must still be used for $x$ below unity.

## X. Inheritance of the Sex-Ratio in the Thoroughbred Racehorse.

By DAVID HERON, M.A.

Mr Cobb in his note in the *Miscellanea* of this number has defined the sex-ratio of an individual as the value which the ratio of male to total births tends to assume when his fraternity becomes indefinitely large.

It is suggested that the variability, $\sigma_0$, actually observed in the sex-ratio as calculated in the ordinary way, arises in part from the variability in what he calls the true sex-ratio, $\sigma_x$, and in part from the error due to considering small families, $\sqrt{\frac{pq}{n}}$, and further that the true correlation between the sex-ratios of parent and offspring can be found by multiplying that obtained from considering the actual families by a factor $(\sigma_x^2 + pq/n)/\sigma_x^2$.

It has already been shewn that $\sigma_x$ is zero, within the limits of probable error, in the case of man. We can however test the reality of this factor directly by two methods. We can compare the correlation found by considering any group of families with that found after rejecting the smaller families, or we can compare the correlation found by considering large completed families with that found from sex-ratios defined by, say, the first four members of these families. In each case, the larger families ought to give higher correlation. The results of the first test have already been given in Prof. Pearson's note. To apply the second test, 1000 mares were taken at random from The General Stud Book, and the sex-ratio of their produce and that of their dam calculated. Only those cases were considered where each mare had at least eight foals, and as only second and subsequent editions of the Stud Book were used, the sex-ratio in each case was based on a *completed* family. From the same families sex-ratios were also calculated from the first four members. Tables I and II give the actual correlation tables

### TABLE I.

*Correlation of Sex-Ratios of Mother and Daughter Mares' Produce: Completed Families. Thoroughbred Horses from The General Stud Book.*

| Sex-Ratios of Mothers' Produce. | Sex-Ratios of Daughters' Produce. | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ·00— ·05 | ·05— ·15 | ·15— ·25 | ·25— ·35 | ·35— ·45 | ·45— ·55 | ·55— ·65 | ·65— ·75 | ·75— ·85 | ·85— ·95 | ·95—1·00 | Totals |
| ·00— ·05 | — | — | — | — | — | — | — | — | — | — | — | — |
| ·05— ·15 | — | — | — | — | 3 | 4 | 2 | 1·5 | ·5 | — | — | 11 |
| ·15— ·25 | — | 1 | 3 | 6 | 9 | 16·5 | 7·5 | 4 | 2 | — | — | 49 |
| ·25— ·35 | 1 | 2 | 7·5 | 14·5 | 33 | 28·5 | 24·5 | 20 | 7 | 2 | — | 140 |
| ·35— ·45 | — | 1 | 10·5 | 28·5 | 56 | 60 | 56 | 29·5 | 14·5 | 2 | 2 | 260 |
| ·45— ·55 | — | 2 | 13 | 33 | 58 | 76·5 | 68·5 | 20·5 | 16·5 | 3 | — | 291 |
| ·55— ·65 | — | 1 | 11 | 14 | 34 | 45 | 31 | 20 | 8 | 1 | — | 165 |
| ·65— ·75 | — | — | 2·5 | 8 | 10 | 23 | 12·5 | 6·5 | 4·5 | — | — | 67 |
| ·75— ·85 | — | — | 1 | 1·5 | 2 | 6 | 4·5 | 1 | — | — | — | 16 |
| ·85— ·95 | — | — | — | — | — | — | 1 | — | — | — | — | 1 |
| ·95—1·00 | — | — | — | — | — | — | — | — | — | — | — | — |
| Totals | 1 | 7 | 48·5 | 105·5 | 205 | 259·5 | 207·5 | 103 | 53 | 8 | 2 | 1000 |

## TABLE II.

*Correlation of Sex-Ratios of First Four Members of Mother and Daughter Mares' Produce. Thoroughbred Horses from The General Stud Book.*

Sex-Ratio of First Four Members of Daughter's Produce.

| Sex-Ratio of First Four Members of Mother's Produce. | $\frac{0}{4}$ | $\frac{1}{4}$ | $\frac{2}{4}$ | $\frac{3}{4}$ | $\frac{4}{4}$ | Totals |
|---|---|---|---|---|---|---|
| $\frac{0}{4}$ | 3 | 12 | 35 | 20 | 10 | 80 |
| $\frac{1}{4}$ | 20 | 65 | 119 | 72 | 15 | 291 |
| $\frac{2}{4}$ | 19 | 104 | 137 | 97 | 23 | 380 |
| $\frac{3}{4}$ | 8 | 45 | 84 | 63 | 13 | 213 |
| $\frac{4}{4}$ | 1 | 8 | 17 | 10 | — | 36 |
| Totals | 51 | 234 | 392 | 262 | 61 | 1000 |

of sex-ratios derived from completed families of eight and upwards and from the first four members of these families respectively, while the means and standard deviations are given in Table III.

## TABLE III.

| Group | | Mean | S. D. |
|---|---|---|---|
| Completed Families ... | Mother's Produce | ·4620 ± ·0029 | ·1379 ± ·0021 |
| | Daughter's Produce | ·5020 ± ·0033 | ·1546 ± ·0023 |
| First Four Members Only | Mother's Produce | ·4585 ± ·0052 | ·2425 ± ·0037 |
| | Daughter's Produce | ·5012 ± ·0052 | ·2426 ± ·0037 |

In the completed family of the mother, at least one female must appear and this of course reduces the mean sex-ratio. When the offspring were not selected in this way, the mean sex-ratio was ·502. If we now apply the condition that at least one member is to be a female, the sex-ratio ought to become $\frac{n-1}{n}$ (·502) where $n$ is the number in the family. The average family was found to be 11·433 and substituting this for $n$ we get the sex-ratio to be ·458. Actually it was found to be ·462, an excellent agreement.

When the question of variability is considered we find as before that $\sigma_x$ is zero. Taking the first four members of the family, the standard deviation is ·2426 ± ·0037, while that which would arise from a chance distribution in families of four is $\sqrt{\frac{·502 \times ·498}{4}} = ·2499$. The difference is within the probable error and it is to be noted that the actual S. D. is less than that which would arise from a chance distribution. According to Mr Cobb's suggestion, it ought to be greater.

If we consider completed families (only the 1000 cases in the second generation were taken) the theoretical variability is given by

$$N\sigma^2 = S\left[n_s\left(\bar{p} - \frac{m_s}{n_s}\right)^2 + pq\,\frac{n_s}{S}\right]$$

in the notation used by Prof. Pearson above, and from this formula $\sigma$ was found to be ·1459, where the term depending on the means contributes practically nothing. Actually the standard deviation calculated from the grouped sex-ratios was found to be ·1546 ± ·0085, so that again the difference is within the limits of the probable error.

Turning now to the question of correlation, in the case of completed families of eight and upwards $r$ was found to be ·0101 ± ·0213, while in the case of the first four members only, $r$ was found to be − ·0064 ± ·0213. These coefficients are both zero within the limits of probable error so that no significant increase in correlation is obtained by raising the number of individuals on which the sex-ratio is based from 4 to an average of 11·4.

It should also be noted that Mr Cobb's corrective factor is *positive* and if the true correlation were at all comparable with its value found for other physical characters, which is of course always positive, the result of our calculation must give a positive correlation. The fact that it gives a small negative correlation is compatible with its being really zero but not with its being the reduced value of a true positive correlation found by using a positive reducing factor.

# NOTICES AND BIBLIOGRAPHY.

## NOTICES*.

*Untersuchungen über das Verhältnis der Kopfmasse zu den Schädelmassen,* by JAN CZEKANOWSKI. Braunschweig, 1907.

An interesting paper giving between 30 and 40 measurements, indices, etc., of each of 119 specimens (65 ♂ and 54 ♀) and the treatment of the observations. The calculations of the mean, median, average deviation, standard deviation, coefficient of variation and coefficient of correlation are explained. The method of moments is referred to, but the $m$th moment is given as $\sqrt[m]{\frac{1}{n}\sum_{\kappa=1}^{n} e_\kappa^m}$, where $e_\kappa$ is the deviation and $n$ the number of cases. An explanation of the measurements is given and various tables giving means etc. follow. Formulæ are given for the calculation of skull from head measurements. Amongst other interesting points the author refers to the variation of the thickness of the skin with the age of the individual at various parts of the head. Czekanowski does not appear to have read any of the very recent English statistical papers and he gives $\cdot 67449 \frac{1-r^2}{\sqrt{n(1+r)}}$ as the probable error of $r$, while in some of his correlations there is probably spurious correlation, e.g. the table on p. 28 which gives the correlation between the cephalic indices of the head and skull measurements.

*Zur Frage der Correlationen der Muskelvarietäten,* by JAN CZEKANOWSKI. New York, 1906 (Boas Memorial Volume).

Certain muscles are not always present, and this paper gives the frequency of the presence of thirteen muscles and the coefficients of correlation between the various muscles. The number of cases investigated is not large and the coefficients vary considerably. The subject would be well worth an extended investigation on the same lines.

*The Numerical Proportions of the Sexes at Birth,* by J. B. NICHOLS. American Anthropological Association, Lancaster, Pa. U.S.A. 1907.

Statistics from Europe, United States, South America, Australia, etc. are used, and the points dealt with are the proportions of births among living and still-born, among legitimate and illegitimate; multiple births; effect of war on sex ratios; differences between ratios in town and country and influence of sanitary environment. A valuable collection of statistics which would have been improved if an estimate of the probable errors had been given in each case.

* Authors of memoirs who desire brief notices of their contents should forward offprints to this Journal, as it is often difficult to procure Journals in which the originals have been issued.

*On the Theory of Correlation for any Number of Variables treated by a New System of Notation*, by G. UDNY YULE. Proceedings of Royal Society, A, Vol. 79, 1907.

A restatement of the theory of correlation. Although the new system of notation does not appeal to one very favourably on a first reading, the results follow easily when the notation has been mastered.

*Frequency Curves and Moments*, by ROBERT HENDERSON. Journal of Institute of Actuaries, July, 1907.

A reprint of a paper which appeared recently in the Transactions of the Actuarial Society of America. It deals with the Sheppard adjustments and Pearson Type curves. The latter are re-numbered and the normal curve appears as Type I, and the Pearson Type III is called II. The paper suffers from the defects of giving in no case a reference to the original author or to a memoir whence the results are reproduced.

*The Scope and Importance to the State of the Science of National Eugenics. The Fourteenth Boyle Lecture*, by KARL PEARSON, F.R.S. London, Henry Frowde, 1907.

If the object of National Eugenics is to be attained it is necessary that the general public, or at least the leaders of public opinion, should appreciate its aims and the results of modern statistical work on heredity and kindred subjects. The general results of the work on eugenics and the practical conclusions to which they tend are dealt with forcibly and clearly in the Lecture, which should do much to awaken interest in the subject among thinking people. We earnestly hope that at no very distant date there will be the "strengthening of racial conscience" and the "scientific basis of conduct" for which Professor Pearson calls.

W. P. E.

## BIBLIOGRAPHY.

(Continued from Vol. v, p. 483.)

(49) ALLEN, B. M. A Statistical Study of the Sex-Cells of *Chrysemys marginata*. Anat. Anz. Bd. xxx, pp. 391—399. 1907.
Counts of the number of sex-cells in embryos of different ages.

(50) Anthropometric Data from Burmah. Calcutta, Office of the Supt. Govt. Print. India, 1906. 235 pp. 8°.

(51) Anthropometric Data from Bombay. *Ibid.* 1907. 341 pp. 8°.

(52) BEAN, R. B. A Preliminary Report on the Measurements of about 1000 Students at Ann Arbor, Michigan. Anat. Record, Baltimore, 1906–7, No. 3, p. 67.
Preliminary report deals mainly with general anthropological characteristics.

(53) BECK, F. R. Eine Methode zur Bestimmung des Schädelinhaltes und Hirngewichtes am Lebenden und ihre Beziehungen zum Kopfumfang. Ztschr. f. Morph. u. Anthropol. Stuttgart, 1906–7, Bd. x, pp. 122—144.

(54) BOAS, F. The Measurement of Variable Quantities. New York, 1906. 52 pp. 8°.
Elementary introduction to subject.

(55) Clark, L. P. & Atwood, C. E. Longevity of Idiots. Medical Record, New York, Aug. 31, 1907.

The authors controvert the popular belief that idiotic children rarely grow to adult life, and as a result of the investigation of 785 cases they conclude that idiocy under good care does not seriously curtail life.

(56) Cook, O. F. Mendelism and other Modes of Descent. Proc. Washington Acad. Sci. Vol. ix, pp. 189—240. 1907.

Critical discussion of Mendelism.

(57) Czekanowski, J. Untersuchungen über das Verhältnis der Kopfmasse zu den Schädelmassen. Arch. f. Anthropol., Brnschwg., 1907, Bd. x, pp. 171—196, 3 pl.

(58) Darbishire, A. D. Some Tables for Illustrating Statistical Correlation. Mem. and Proc. Manchester Lit. and Phil. Soc. Vol. li, Part iii, 21 pp. 1907.

Popular discussion of the genesis of correlation on the basis of chance.

(59) Darrach, W. Variation in the Postcava and its Tributaries as observed in 605 examples of the Domestic Cat. Amer. Jour. Anat. Vol. vi, No. 3 (Proc. Assoc. Amer. Anat.), pp. 30—33. 1907.

(60) Donaldson, H. H. A Comparison of the White Rat with Man in Respect to the Growth of the Entire Body. Boas Memorial Volume, New York, 1906, pp. 5—26, 1 pl.

Follows growth of rat in body weight from conception to maturity. Growth curves exhibit similar phases in rat and man, and the growth of the ♂ in relation to that of the ♀ is similar in the two cases.

(61) Elderton, Ethel M. Assisted by K. Pearson. On the Measure of the Resemblance of First Cousins. Eugenics Laboratory Publications, No. iv. (Dulau & Co., London.)

Deals with the degree of resemblance of first cousins for physical, psychical and pathological characters.

(62) Fischer, E. Die Bestimmung der menschlichen Haarfarben. Korrespondenzblatt der deut. Anthropolog. Gesellschaft, Jg. xxxviii.

A lengthy discussion without fresh chemical or microscopic analysis of human hair tints. Fischer makes two series which have a fundamental difference in tone, the one being grey black and the other yellow brown. He condemns the scale published in this Journal—admittedly somewhat unnatural in tints—because the grey shades are absent. Only one such grey shade occurred in the samples collected, and it was probably of Slavonic origin. Fischer's own scale is prepared from a "künstlicher Glanzstoff aus Zellulose"; it remains to be seen whether the dye on these threads will be permanent and capable of accurate reproduction. The author says it may be purchased for 20/- from Franz Rosset in Freiburg i. B., but we have been unable so far to procure a copy.

(63) Forbes, S. A. An Ornithological Cross-Section of Illinois in Autumn. Bull. Ill. State Lab. Nat. Hist. Vol. vii, Art. 9, pp. 305—335. 1907.

An attempt at an exact determination of the relative frequency of different species of wild birds and the relation of these frequencies to various environmental conditions, in a strip of land 150 feet wide extending across the State (total area covered = 3519 acres). Interesting detailed statistical tables showing the results of this ably conceived and executed animal census are given.

(64) Hatai, S. Biometrical Studies on the Skull of the Albino Rat. Anat. Record, Baltimore 1906–7, No. 3, p. 51.

A preliminary report.

(65) HEBERT, P. Z., M.D. Killing of the Unfit, and the Transmissibility of Acquired Characters. General Practitioner, July 6, Nov. 16 and 23, 1907.

Of no interest to the biometrician or to the biologist.

(66) HERMS, W. B. An Ecological and Experimental Study of Sarcophagidae with Relation to Lake Beach Debris. Jour. Exper. Zool. Vol. IV, pp. 45—83. 1907.

Contains, *inter alia*, some interesting data on growth in beetles.

(67) HERON, D. A First Study of the Statistics of Insanity and the Inheritance of the Insane Diathesis. Eugenics Laboratory Publications, No. II. (Dulau & Co., London.)

(68) HOFMANN, KARL. Der exakte Artbegriff, seine Ableitung und Anwendung. Annalen der Naturphilosophie, Bd. VI, S. 154—216.

The author proceeds from an impossible definition of a species due to Heincke (which involves the non-existence of Quetelet's "mean man"!) to develop a theory of species by which each species is represented by an $n$-dimensional sphere. We need only say that a study of actual variation in multiple correlation would have convinced the writer that Heincke's and his own fundamental hypothesis is invalid.

(69) HOOKER, R. H. Correlation of the Weather and Crops. Jour. Roy. Stat. Soc. Vol. LXX, Part 1. 1907.

(70) HRDLICKA, A. Measurements of the Cranial Fossae. Proc. U. S. Nat. Mus., Washington, 1907, pp. 177—232.

"This paper deals with the absolute and relative lengths of the cerebral and cerebellar fossae in man and a series of animals, and with the relation of the length of the different fossae to the form of the skull."

(71) KELLICOTT, W. E. Correlation and Variation in internal and external characters in the Common Toad (*Bufo Lentiginosus Americanus*, Le C.). Journal of Experimental Zoology, IV, pp. 575—614.

An interesting paper with one or two slight misunderstandings. Thus there is no difficulty whatever (as is suggested on p. 576) in calculating multiple correlations for 3 (or even 4) variables when the coefficients for the correlation of every pair are known. Again in reference to the remark on p. 610, we may say that Pearson starts his paper by excluding the influence of growth, and the results cited from Greenwood, p. 611, are largely due to correlated pathological changes, one phase of growth. The fact that Kellicott does not take age into account ("the subjects were of various ages") renders his correlations largely 'spurious'; it might possibly account for the markedly higher correlations of the females and some light on this point might be obtained by correlating indices. Finally the belief that the *whole* of a small population, said to be taken in this case, is better than a random sample of a very much larger population is not well founded, for the total population in such a case may be due to comparatively few progenitors.

(72) KERR, A. T. Statistical Studies of the Brachial Plexus in Man. Amer. Jour. Anat. (Proc. Assoc. Amer. Anat.), Vol. VI, No. 3, pp. 53, 54. 1907.

(73) KIAER, A. N. Rapport sur la statistique de la fécondité du mariage. Bull. de l'Instit. internat. de statist. Londres, 1906, T. XV, livr. 2, pp. 398—401.

(74) KING, HELEN DEAN. Food as a Factor in the Determination of Sex in Amphibians. Biol. Bull. Vol. XIII, pp. 40—56. 1907.

Shows that the sex-ratio is not significantly influenced by the character of the food in *Bufo*.

(75) KIRKOFF, N. Recherches anthropologiques sur la croissance des élèves de l'École militaire de S. A. R. le Prince de Bulgarie, à Sofia. Bull. et Mém. Soc. d'Anthrop. de Paris, 1906, Sér. 5, T. VII, pp. 226—233.

(76) KLEBS, G. Studien über Variation. Arch. f. Entwickelungsmech. Bd. 24, pp. 29—113. 1907.

An extensive statistical study of variation in plants, dealing mainly with *Sedum spectabile* and having as its object the investigation of the "notwendigen Zusammenhang der Variationskurven mit bestimmten Aussenbedingungen."

(77) LAPICQUE, L. & GIRARD, P. Sur le poids de l'encéphale chez les animaux domestiques. C.R. Soc. Biol., T. 62, No. 19, pp. 1015—1018. 1907.

(78) LEICESTER, J. C. H. The relative Sizes of the Maternal Pelvis and of the Foetus in Europeans, Eurasians, East Indians and Bengals. Lancet, 1907, Vol. I, pp. 150—153.

(79) LOCK, R. H. On the Inheritance of Certain Invisible Characters in Peas. Proc. Roy. Soc. Vol. 79B, pp. 28—34. 1907.

(80) MACCURDY, H. & CASTLE, W. E. Selection and Cross-breeding in Relation to the Inheritance of Coat-pigments and Coat-patterns in Rats and Guinea-pigs. Carnegie Institution of Washington, Publication No. 70, pp. iii and 50. 1907.

"In both rats and guinea-pigs we can by selection increase or decrease at will the average extent of the pigmented areas. In both rats and guinea-pigs the extent of the pigmented areas varies continuously, and out of these continuous variations permanent modification of the pigmentation can be secured"....."Further we are far from convinced that all evolutionary progress is to be attributed to discontinuous variation any more than to Mendelian inheritance."

(81) MARTIN, R. System der (physischen) Anthropologie und anthropologische Bibliographie Korrespondenzblatt der deutschen Anthropolog. Gesellschaft, Jg. XXXVIII.

A most elaborate and most German numerical symbolism for the classification of the subject matter of Anthropology. The title of Martin's own paper is $\begin{cases} 573\ (01) \\ 573\ (091) \end{cases}$ in this symbolism, which we suppose gives some satisfaction to the classificatory mind.

(82) MICHAELIS, F. Das Hirngewicht des Kindes. Monatsschr. f. Kinderkrankh., 1907–8, Bd. VI, S. 9—26.

(83) NETTLESHIP, E. A History of congenital stationary Night-Blindness in nine consecutive generations. Ophthalmological Society's Transactions, Vol. XXVII.

A fine piece of work, if we do not wholly agree with the author's interpretation of his statistics.

(84) NEUSTATTER, O. Geburtenziffer und Fruchtbarkeit. München. med. Wochenschr., Bd. LIV, 85 pp. 1907.

(85) NICHOLS, J. B. The Numerical Proportion of the Sexes at Birth. Mem. Anthropol. Assoc. Vol. I, pp. 249—300. 1907.

(86) PEARL, R. & CLAWSON, A. B. Variation and Correlation in the Crayfish, with special Reference to the Influence of Differentiation and Homology of Parts. Carnegie Institution of Washington, Publication No. 64, pp. 70. 1907.

(87) PEARSON, KARL. Mathematical Contributions to the Theory of Evolution. XVI. On Further Methods of determining Correlation. (Dulau & Co., London.)

Gives methods of finding true variate correlation from placing individual in rank.

(88) ROMMEL, G. M. Relative Proportion of the Sexes in Litters of Pigs. U. S. Dept. Agr. Bur. Anim. Ind. Circular No. 112, p. 1. 1907.

Finds that out of 13,285 pigs born in 1477 litters there were 6660 ♂♂ and 6625 ♀♀, giving the sex ratio 1005 ♂ to 1000 ♀. Eight breeds were represented in the lot.

(89) ROMMEL, G. M. & PHILLIPS, E. F. Inheritance in the Female Line of Size of Litter in Poland China Sows. Proc. Amer. Phil. Soc., Vol. XLV, pp. 245—254. 1906.
An abstract of this paper has appeared in Biometrika, Vol. v, p. 203.

(90) SAMTER, M. Das Messen toter und lebender Fische für systematische und biologische Untersuchungen. Arch. f. Hydrobiol. u. Planktonkunde, Bd. II, pp. 143—185. 1906.
Describes a photographic method of measuring soft-bodied organisms, likely to be useful in many cases.

(91) SCHUSTER, E. The Promise of Youth and the Performance of Manhood. Eugenics Laboratory Publications, No. III. (Dulau & Co., London.)

(92) SHULL, G. H. Elementary Species and Hybrids of Bursa. Science, N. S., Vol. XXV, pp. 590, 591. 1907.
Preliminary discussion of the results obtained from rearing "over 20,000 pedigreed specimens of *Bursa bursa-pastoris* (L.) Britton."

(93) SPILLMAN, W. J. Colour Inheritance in Mammals. Science, N. S., Vol. XXV, p. 313. 1907.

(94) ——. Inheritance of the Belt in Hampshire Swine. *Ibid.* Vol. XXV, pp. 541—543. 1907.

(95) SRDÍNKO, OTAKAR. Poměr pohlavi při porodech v rakousku vůbec a v čechách zvláště. [The Sex Ratio of Births in Austria and especially in Bohemia.] Předbežné sdělení, z Časopis lek. česk., Ročnik, 1907, 51 pp.

(96) TOYAMA, K. Studies on the Hybridology of Insects. I. On some Silk-worm Crosses, with special Reference to Mendel's Laws of Heredity. Bull. Coll. Agric., Tokyo Imp. Univ., Vol. VII, pp. 259—393. 1906.
Embodies an extensive mass of detailed data on the results of hybridizing silk-worms. Results partly Mendelian, partly not.

(97) VIGOR, H. D. & YULE, G. U. On the Sex-Ratios of Births in the Registration Districts of England and Wales, 1881-90. Jour. R. Stat. Soc., Vol. LXIX, Part iii. 1906.

(98) VOGT, H. Studien über das Hirngewicht der Idioten; das absolute Gewicht. Monatsschrift f. Psychiat. u. Neurol., Bd. XX, pp. 424—469. 1906.

(99) WEINBERG, W. Verwandtenehe und Geisteskrankheit. Arch. Rassen. u. Gesellsch. Biol. Jahrg. 4, pp. 471—475. 1907.
Discussion of Mayer's statistics (Jahrb. internat. Ver. f. Rechtswiss. u. Volkswirtsch., Bd. VI—VII, 1904).

(100) WHIPPLE, G. M. A Quick Method for Determining the Index of Correlation. Amer. Jour. Psychol., Vol. XVIII, pp. 322—325. 1907.

(101) WRIGHT, A. H. A Graphic Method of Correlating Fish Environment and Distribution. Amer. Nat., Vol. XLI, pp. 351—354. 1907.

(102) YULE, G. U. On the Theory of Correlation for any Number of Variables, treated by a New System of Notation. Proc. Roy. Soc., Vol. 79A, pp. 182—193. 1907.